FORCE AND FLUCTUATIONS IN POLYMER ENSEMBLES

A Dissertation
Presented to
The Academic Faculty

By

James T. Waters

In Partial Fulfillment
of the Requirements for the Degree
Doctor of Philosophy in the
School of Physics

Georgia Institute of Technology

December 2016

FORCE AND FLUCTUATIONS IN POLYMER ENSEMBLES

Approved by:

Dr. Harold D. Kim, Advisor
School of Physics
Georgia Institute of Technology

Dr. Jennifer E. Curtis
School of Physics
Georgia Institute of Technology

Dr. James C. Gumbart
School of Physics
Georgia Institute of Technology

Dr. Stephen C. Harvey
Perelman School of Medicine
University of Pennsylvania

Dr. Kurt A. Wiesenfeld
School of Physics
Georgia Institute of Technology

Date Approved: October 12, 2016

For my father

ACKNOWLEDGEMENTS

I would like thank the other members of the Harold Kim Lab, past and present. To Bo, Gable, Jiyoun, Rasesh, Tae and Tung, I would like to say I value you not only for their input and feedback on research, but for providing the human contact necessary to preserve my sanity (such as it is). I would also like to express my appreciation to Dr. Lauren Pawlak, for her patience as I slowly make my way to the finish line.

I am grateful to my committee members for their time reading and considering the work represented here. Additionally, I appreciate guidance from James Gumbart and and Stephen Harvey as I was learning new skills in molecular dynamics simulation. While that work falls outside the scope of this thesis, I feel it has been an important part of my development as a scientist. I also thank Kurt Wiesenfeld for useful discussions which offered a fresh perspective when I was struggling or stuck.

Finally, I would like to thank Harold Kim for the opportunity to work in his lab. I recognize that coming to him as theorist with no background in biology or in experiments may have seemed a strange proposition four years ago, and I appreciate him taking a chance on me. I hope that the scientific work that resulted from this collaboration will prove useful towards the research program of the Kim Lab after I have gone, and that the methodologies developed in conducting this work will find application even beyond my own or Harold's purview.

TABLE OF CONTENTS

LIST OF FIGURES

SUMMARY

The focus of this work is on developing mechanical models for studying polymer chains under different constraints and confining geometries. This will have application to biological processes, where the bending and looping of DNA molecules plays an important role in transcriptional regulation, replication, repair, and packaging. By using a general model, we produce results which may prove relevant for a range of polymeric problems.

Our model attempts to bridge problems solvable with classical mechanical and statistical mechanical methods. Average forces in different geometries of polymer confinement and constraint can be found from thermodynamic arguments, however by treating the problem as a classical mechanical system we can map the phase space onto a whole distribution of forces. This will reproduce the average result, but also provide new information pertinent to kinetics and lifetimes governed by fluctuations away from the mean behavior.

Confinement takes the form of a hard bounding surface, restricing the conformational space accessible to a long chain molecule. Entropically, the chain will try to break away from this confinement, and there is a corresponding change in the distribution of mechanical forces. Constraint takes the form of a fixed distance between the ends of the chain, with an average constraint force corresponding to the thermodynamic mean force. We find in both instances that the distribution of forces is aysmmetric and non-Gaussian, implying that the average force is not the most probable force. It also implies that large outliers may appear more frequently than a normal distribution would predict.

CHAPTER 1
INTRODUCTION

1.1 Investigating the Mechanical Properties of DNA

DNA is one of the most important molecules in nature. By encoding genetic instructions for proteins, it serves as a cookbook for all life on the planet. Though the mapping from codons to amino acids is known, many unsolved problems exist in understanding how the transcription of this information is regulated. This regulation is fundamental to generating the complexity of life observed in the world today, and will depend not just on what the sequence codes for but on the mechanical properties of the DNA molecule- how it behaves as an object in space under the laws of physics.

Motivated by a desire to explore these properties, we began investigating the specific effects of confinement on DNA bending and looping. Bending of DNA plays many important biological roles. DNA bound in loops by proteins can control access to promoter regions [1]. Bending plays a role in DNA replication[2, 3], in DNA repair, or in the integration of one DNA sequence into another[4]. Tight bending of DNA around proteins also plays an important role in how genomes are packaged within the nucleus[5]. Additionally, DNA has proven to be a useful building block for self-assembling structures, such as rings[6] and cubes[7]. DNA has also been used for constructing molecular-scale motors[8, 9]. It is also commonly employed as a 'handle' for unfolding RNA molecules[10], transmitting force applied via an optical trap to the molecule under study. Expanding our working knowledge of DNA as a material, and not simply a carrier of information, can aid in the design and predicting the stability of such structures.

To probe forces in loops and other conformations relevant for biological function, we turn to a coarse-grained model of the DNA strand as a semi-flexible polymer chain. Al-

though these systems can be explored in great detail with molecular dynamics simulations, such treatments present challenges of size and timescale. The dynamics of loop breaking, for instance, may occur on the order of minutes[11], beyond what is tractable with all-atom simulations. Instead, a coarse-grained approach combining theory and computation offers a level of detail beyond analytically soluble thermodynamics, but tractable on scales larger than those accessible by molecular dynamics[12]. Furthermore, as this coarse representation of DNA omits specific details of individual base pairs, the results produced here apply generally to other polymer chains. Other important biological properties may be governed by fluctuations in forces[13, 14], and so by developing a non-specific model, we can find results that will apply to a wide range of systems.

In this work, we apply new methods to existing coarse-grained models to sample polymer states in both position and momentum space. Choosing points in the full phase-space, as opposed to the conformational space alone, allows us to compute the distribution of constraint forces within this coarse model. Equilibrium statistical mechanics alone can give us the average force associated with phenomena such as loop-breaking or surface desorption. The rate of such processes, though, may be influenced by the range of fluctuation in this force and not simply its mean value [15]. Knowing this distribution in turn can inform a detailed picture of the kinetics as DNA strands undergo conformational changes, using information inaccessible from equilibrium statistical mechanics.

1.2 Entropic Force

Confining or constraining a DNA strand, or another long chain molecule, will restrict the number of conformations available to it, and as a result the system will be subject to an entropic force. The tendency of a system to move towards higher entropy is described by the second law of thermodynamics. While entropy is popularly understood to be synonymous with disorder, it is more precisely a measure of the number of states or configurations accessible to the system. This resulted in the celebrated entropy formula given by Boltz-

mann[16]

$$S = k_B \log W \tag{1.1}$$

For a flexible polymer, the set of states is the number of possible chain conformations. For an inextensible chain, this is analagous to the number of random walks of a specified length. In one dimension, the distribution of possible final positions relative to the starting position is given exactly by the binomial distribution. However, it is more easily approximated as a Gaussian distribution when the number of steps grows large.

$$P(r) = \sqrt{\frac{1}{2\pi N a^2}} \exp(-r^2/2Na^2) \tag{1.2}$$

In three dimensions, the probability along any given axis will be given by the same distribution. For the scalar extension, though, we find that each possible extension is weighted by a geometric factor corresponding to the surface area of a spherical shell. This gives us a probability distribution in three dimensions

$$P(r) = \sqrt{\left(\frac{3}{2\pi N a^2}\right)^3} 4\pi r^2 \exp(-3r^2/2Na^2) \tag{1.3}$$

which has a local maximum at $r = \sqrt{2Na^2/3}$. Conformations at extensions above and below this most probable value are expected to move inward or outward, respectively, over time. Correspondingly, there is a net average force

$$F(r) = k_B T \left(\frac{2}{r} - \frac{3r}{Na^2}\right) \tag{1.4}$$

There is no fundamental force specific to this progression towards states of higher entropy- rather, it arises from the interactions of the polymer chain with the heat bath. Therefore we expect this average force to emerge from inertia within the flexible chain as it moves about in a heat bath. To understand this in greater detail, we must first coarse-grain the polymer chain into a series of point masses which we can study using classical mechanics.

1.3 The Kratky-Porod Model

1.3.1 The Continuous Wormlike Chain

The most common simplification of DNA or other long polymers is the worm-like chain (WLC) model[17]. This model omits the complex structural details of polymeric units, instead capturing the intrinsic stiffness of the molecule with a bending potential. In such a model the shape of a molecule is represented by a one-dimensional curve in space of a given contour length L. This can be described by a vector function $\mathbf{r}(s)$ in three-dimensional space. The tangent vector can be found by differentiating this function.

$$\hat{t}(s) = \frac{\partial \mathbf{r}(s)}{\partial s} \tag{1.5}$$

Every possible function satisfying the length requirement is accessible, but a potential energy associated with bending the chain reduces the probability of sharply curved conformations. This energy is assumed to be quadratic in the curvature[18], defined in terms of the tangent vector.

$$U = \frac{L_P}{2} \int_0^L \left(\frac{\partial \hat{t}(s)}{\partial s} \right)^2 ds \tag{1.6}$$

This bending energy is parameterized in terms of a persistence length L_P, which determines the scale over which orientational order is preserved in the molecule. The expected dot product between tangent vectors at different points along the chain will fall of exponentially with separation, at a rate proportional to this persistence length.

$$\langle \hat{t}(s) \cdot \hat{t}(s + \Delta) \rangle = \exp(-\Delta / L_P) \tag{1.7}$$

The distribution of contours can be summarized by statistics of the extension $\mathbf{R}$, or end-to-end distance. This is equivalent to the sum of all tangent vectors

$$\mathbf{R} = \int_0^L \hat{t}(s)ds = \mathbf{r}(L) - \mathbf{r}(0) \tag{1.8}$$

A perfectly straight conformation, corresponding to the extension $|\mathbf{R}| = L$, will have no energy cost, but will be unlikely due to the extremely low entropy of the state. Conversely, tightly coiled conformations are numerous and therefore entropically preferable, but their large associated bending energy imply a low probability of the molecule being in such a state. The end result will be a distribution of end-to-end distance peaked between minimal and maximal extension. The location and width of this peak will depend on the length of the contour relative to the persistence length. The simplest way to characterize the distribution is in terms of the mean-squared extension.

$$\langle R^2 \rangle = \langle \mathbf{R} \cdot \mathbf{R} \rangle = 2L_P L \left(1 - \frac{L_P}{L}(1 - e^{-L/L_P}) \right) \tag{1.9}$$

In the limit of small L/L_P, we have $\langle R^2 \rangle \approx L^2$, describing a rod-like conformation. In the limit of large L/L_P, we have $\langle R^2 \rangle \approx 2L_P L$, describing a random walk of $L/2L_P$ steps, each of length $2L_P$. This is a way of describing the intuition about energetic versus entropic trade-offs above in more quantitative terms.

1.3.2 Measuring Chain Parameters

The relationship between extension and flexibility described above allows us to infer the persistence length of physical polymers from their conformational statistics. DNA persistence length can be measured experimentally by rotational diffusion[19], cryo-electron microscopy[20] and force-extension measurements[21]. These observations have determined a persistence length of 45-53 nm, corresponding to roughly 150 base pairs. DNA pulling experiments on long strands have shown agreement with a wormlike chain in the flexible

limit described above[22, 23]. Cyclization experiments on shorter strands have also shown agreement with ring-closure probabilities for the wormlike chain derived by Yamakawa[24, 25]. Representation of the chain as inextensible is justified by the relative stiffness of the DNA molecule along this direction[26].

Additonal details of the polymer chain, such as excluded volume interactions, are neglected. For expediency, we instead treat the wormlike chain as a 'ghost' chain, able to pass through itself. In addition to being computationally quicker, this avoids the rare possibility of getting stuck in knotted conformations when perturbing the chain coordinates. For much of this work, the polymers under consideration will be of the order of a few persistence lengths or less. On this scale, collisions within the chain are rare. Excluded volume interactions have been shown to not have an observable effect on the size of a conformation until the contour length is over an order of magnitude larger than the persistence length[27].

1.3.3 Discretization and Coordinate Space

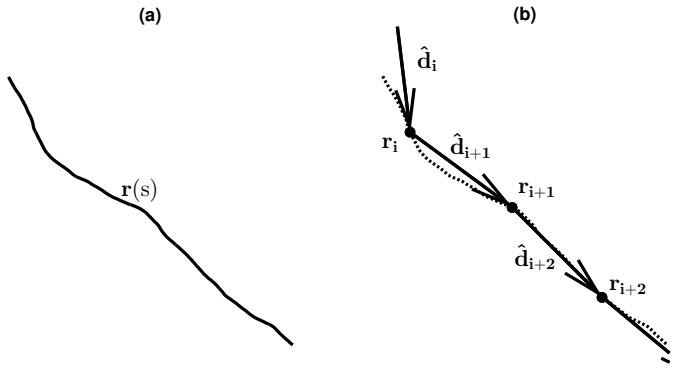

Figure 1.1: Alternate representations of a chain conformation. A wormlike chain is represented by a continuous, differentiable curve. The Kratky-Porod model represents the chain with a set of bonds and point masses.

The infinite-dimensional functional space of the wormlike chain model is convenient for path-integral treatments of polymer distributions[28, 29], and several functions of the

distribution can be calculated analytically. Once these calculations become more complex, however, we must turn to computational means. Simulation, in such an event, will be more easily accomplished if the chain is represented by a set of discrete elements. The Kratky-Porod model[17] represents a decomposition of a continuous wormlike chain into a series of mass points connected by inextensible bonds. The entire strand is represented with a set of N vectors of equal length l. The bending energy is represented by a harmonic function of the angle between successive bonds.

$$U = \frac{L_P}{2l} \sum_{i=1}^{N-1} \cos^{-1} \left(\frac{\mathbf{d}_i \cdot \mathbf{d}_{i+1}}{l^2} \right)^2 \tag{1.10}$$

In the limit of a short segment length, this will converge on the distribution for the continuous wormlike chain.

In the model described above, N links will correspond to $N + 1$ mass points in three-dimensional space, giving $3N + 3$ cartesian coordinates. However, the N inextensibility constraints of the polymer chain will constrain us to a $2N + 3$ dimensional manifold within this space. This may be further reduced by fixing the end-to-end distance, or the position of the first monomer. Tension along the chain can be computed from the force conjugate to the constrained coordinates. A single particle tethered a fixed distance from the origin offers a simple example of such a manifold (Figure 1.2), where it explores the two-dimensional surface within a three-dimensional cartesian space. The force of constraint within the tether will vary with the particle's momentum. This then allows us to turn a distribution of momenta and position into a distribution of force. Typically, the distribution of momenta will be chosen assuming the particle is in a heat bath at some temperature. For a particular set of coordinates, such as the azimuthal and altitudinal coordinates shown in the figure, the precise relationship between momentum coordinates and force may vary also with position in space, as one value of p_ϕ may mean more or less motion further or closer to the equator. This is perhaps the simplest example of how kinetic energy at one point in momentum

varies with how the manifold is curved.

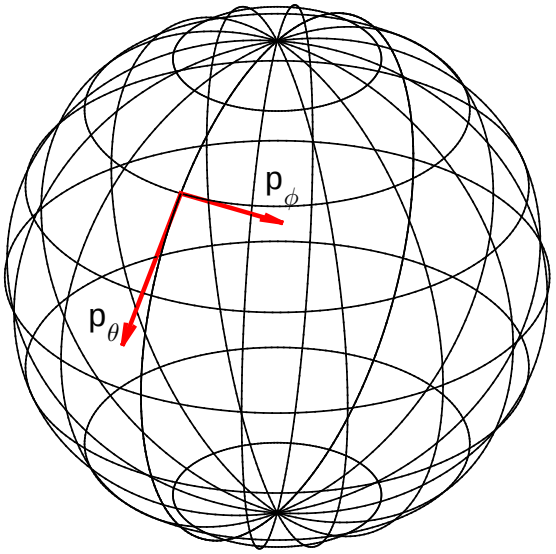

Figure 1.2: The sphere represents a two-dimensional manifold in three-dimensional space. The kinetic energy associated with a particular momentum value varies with position on the sphere.

The curvature of the manifold can be found from the Jacobian, which describes how the $3N+3$ cartesian coordinates change with respect to the $2N+3$ free coordinates.

$$J = \begin{pmatrix} \frac{\partial \mathbf{r}_1}{\partial q^1} & \cdots & \frac{\partial \mathbf{r}_1}{\partial q^{2N+3}} \\ \vdots & \ddots & \vdots \\ \frac{\partial \mathbf{r}_N}{\partial q^1} & \cdots & \frac{\partial \mathbf{r}_N}{\partial q^{2N+3}} \end{pmatrix} \tag{1.11}$$

This relationship between cartesian and free coordinates defines a metric tensor, $M = J^T J$, which tells us everything we need to know about the shape of the manifold. The kinetic energy is defined in terms of this tensor.

$$T = \frac{1}{2}\dot{q}^i M_{ij} \dot{q}^j \text{ where } M_{ij} = \sum_{k=1}^{N} \frac{\partial \mathbf{r}_k}{\partial q^i} \cdot \frac{\partial \mathbf{r}_k}{\partial q^i} \tag{1.12}$$

The spatial coordinates q^i and their derivatives are written with superscripts, denoting their status as contravariant vectors that vary inversely with the basis vectors in our coordinate

system. Momenta are represented as covariant vector components p_i, and any index repeated as a subscript and superscript is summed over. At every point, we can define the tensor M^{ij} as the inverse of the covariant metric, such that $M_{ij}M^{jk} = \delta_i^k$.

For any given values of the coordinates q^i, this allows us to find the size of the corresponding momentum space by integrating over the momentum coordinates. Weighting each state by its kinetic energy, the volume element is given by

$$\int_{-\infty}^{\infty} \exp\left(-\frac{\beta}{2} p_i M^{ij} p_j\right) d^n p = \left(\frac{2\pi}{\beta}\right)^{n/2} \sqrt{|M|} \tag{1.13}$$

where the absolute value of a matrix quantity denotes a determinant and n is the number of coordinates. This can be incorporated into the weighting of different states in position space when computing the partition function, though the numerical prefactor $(2\pi/\beta)^n$ can be omitted, as it will be the same for all values of the spatial coordinates. For the example of the spherical surface, M can be written as $r^2 d\theta^2 + r^2 \sin^2\theta d\phi^2$, and this returns the familiar surface element $r^2 \sin\theta$.

While the base pairs of DNA may define a natural scale at which to apply this discretization, ultimately it is an arbitrary decision. This amibguity means that if we want to be confident that our simulation results have physical relevance, they should be independent of this parameter. However, this arbitrariness can have the advantage that long DNA strands can be represented by a number of units less than their number of base pairs, each with a longer length. Altering the scale simply requires adjustment of the bending energy at each pivot point to preserve the overall persistence length. This will also enable us to generalize our results from the specific application of DNA to the more general case of an arbitrary semiflexible chain.

1.3.4 Types of Confinement

We consider two distinct ways of limiting the coordinate space the discretized chain can explore. A holonomic constraint can hold one parameter of the system fixed at a specific value, reducing the degrees of freedom for the system, as in Figure 1.3(a). The specific example shown in the figure corresponds to a two-monomer chain (dimer) confined to two dimensions, with extension r held at a fixed value. Using generalized coordinates q^1 and q^2 which range from 0 to 2π, representing the orientation of the two bonds, Figure 1.3(b) shows how this will confine us to one of the red contours defined as $r = 2\cos\left(\frac{q^1-q^2}{2}\right)$. Each value of r selects a different contour.

Alternatively, a bounding surface will leave the same number of degrees of freedom, but will prevent the system from exploring their full range, as in Figure 1.3(c). This corresponds to the same two-dimensional dimer, but with a restriction that the end-points of the bonds cannot cross the surface $y = -h$. Changing h changes the range of accessible angles, as in Figure 1.3(d). In both events, the change in the partition function corresponds to a change in free energy and an attendant force.

The canonical example of the first case is the entropic force of a freely-jointed polymer at fixed extension, used to explain the elasticity of rubber-like materials[30, 31]. This will reduce the number of coordinates in the partition function integral, Equation 1.16, selecting only the states at a specific value of the constrained coordinate q^i. The velocity corresponding to this coordinate ($\dot{q}^i$) is assumed to be fixed at zero.

$$\begin{aligned} Z(\xi) &= \int\int \exp(-\beta\mathcal{H}(q,p))\delta(q^i-\xi)\delta(\dot{q}^i)d^Nqd^Np \\ &= \int\int \exp(-\beta\mathcal{H}'(q,p;\xi))d^{N-1}qd^{N-1}p \end{aligned} \tag{1.14}$$

Contributions will come from the volume of space increasing or decreasing with ξ, as well as changes with the average energy along this coordinate.

In the second case, the effect is not in the integrand for the partition function but on the

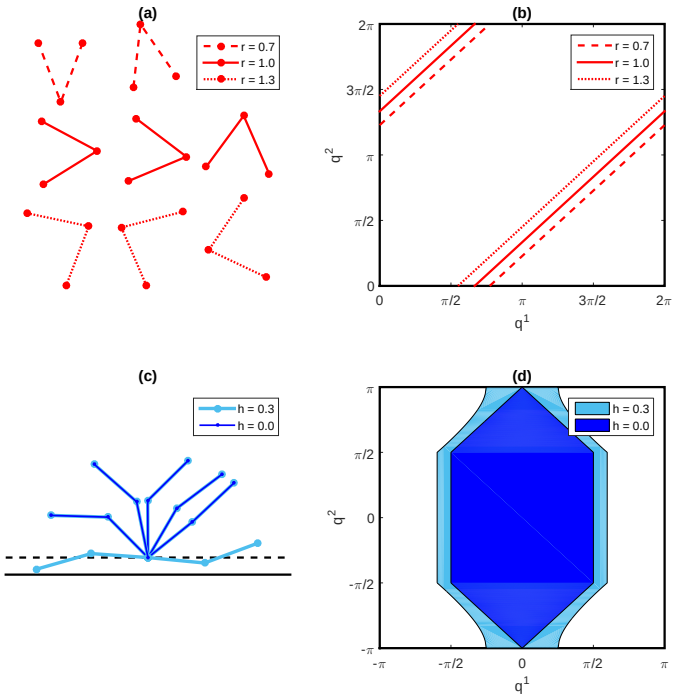

Figure 1.3: (a,b) Holonomic constraints confine the system to lower dimensional subspace. Changing the value of the constraint parameter moves the system to an adjacent subspace. (c,d) Boundaries restrict the space but do not change the dimensionality. Moving the boundary increases or decreases the accessible range of spatial coordinates.

bounds of integration.

$$Z(\xi) = \int \int_{q_-(\xi)}^{q_+(\xi)} \exp(-\beta H(q,p)) d^N q d^N p \tag{1.15}$$

Note that the limits on the momentum integral are the same for all values of ξ. The change in partition function will be positive or negative depending on whether more states are added or removed as the boundary coordinate changes. This same reasoning can allow us to derive the ideal gas law for particles confined to box[32].

1.4 Mean Forces and Fluctuations

After establishing the particulars of our coarse-grained model, we can now investigate how it behaves in an environment. We assume the polymer chain exists in a bath of some solvent, at a thermodynamic temperature $k_B T$. Some of the forces and dynamics we are interested in can be gleaned from a purely statistical mechanical treatment, while others will require a fully classical mechanical approach. We begin with the first of these.

1.4.1 Partition Function and Free Energy

The canonical partition function contains a complete picture of our system. This is derived from an integral over all possible values of position and momentum coordinates, with each microstate weighted by a Boltzmann factor corresponding to its energy

$$Z = \int \int \exp(-\beta \mathcal{H}(q,p)) d^N q d^N p \tag{1.16}$$

where β is the inverse temperature $1/k_B T$. The Hamiltonian represents the total energy of the chain. This is a sum of the potential energy U, which depends only on the spatial coordinates q^i, and the kinetic energy T. In cartesian space, kinetic energy is independent of position, varying only with the momentum coordinates p_i. Due to the fact that our system is confined to a curved manifold, this is not guaranteed, and instead we rely on the metric tensor to describe how this quantity varies in phase space. The Helmholtz free energy is derivable from the partition function

$$A = -k_B T \log(Z) \tag{1.17}$$

and corresponds to the potential of mean force.

$$\langle f \rangle = \frac{\partial A}{\partial \xi} \tag{1.18}$$

This means that by quantifying how the partition function varies with a constrained coordinate, we can obtain an average value of the force required to keep the system constrained.

1.4.2 Sampling Force Distributions

Although the probability of occupying different states can be found from their respective free energies, obtained from the partition function, this does not describe the dynamics of how rapidly the system may shift between the states. Theory predicts that the rates of biomolecular reactions may be influenced not only by average forces, but by the distribution of forces a molecule experiences over time [15]. Kinetics of processes like DNA unwinding from nucleosomes are expected to be influenced by fluctuating forces[33]. Non-exponential kinetics observed in protein unfolding[34] may also arise from processes not dependent on the average force alone. Probing the distribution of forces is one avenue that may yield insight. Intuitively, we expect parameters corresponding to high force fluctuations are associated with less stable states and parameters corresponding to smaller fluctuations correspond to more stable states.

While obtaining the average force is straight-forward, deriving the range of fluctuations from the partition function is more difficult. Instead, we turn to a Monte Carlo sampling method to obtain an ensemble of force values. By introducing a Lagrangian function describing the system, we can compute an undetermined multiplier conjugate to the constrained coordinate for any value of the position and momentum variables. Computing this force requires not only a set of sampled conformations, but sampled momentum coordinates as well. These can be obtained by assuming equipartition. For a set of independent coordinates, equilibrium momenta can be sampled from independent Gaussian distributions. For coordinates that are coupled together by a metric tensor, we can assume an alternative set of coordinates that is complicated to define geometrically but which correspond to a diagonal metric tensor[35]. By applying equipartition in these coordinates, and converting back to our original coordinates, we obtain a realistic sampling of momenta which can be used to

find a sampling of forces.

The classical mechanical constraint forces, found by sampling the ensemble of states in both position and momentum space, are found to agree on average with the mean force predicted by the free energy. This agreement provides a mechanistic explanation for behavior typically understood in a purely thermodynamical sense. After establishing the connection between the partition function and mean mechanical force, we can also employ this process in reverse[36, 37]. Rather than using a derivative of a known free energy obtained from a known partition function to verify our sampled mean constraint force, we can find free energy derivatives in cases where the partition function is unknown.

Our sampling of momentum space assumes that at any given instant the velocities correspond to an equilibrium distribution. In reality, the reactions under consideration may be taking place in a crowded cellular environment. This environment implies an overdamped system, where the molecules in question are treated as inertia-less. The result is that accelerations of coordinates are dominated by random terms associated with molecule-solvent interactions. The rate at which velocities, and thereby generalized forces, change in time, may also impact kinetics [15], but it is not the focus of the work here.

1.5 Structure of the Thesis

The second chapter of this thesis will describe the effects of confinement on the equilibrium properties of polymer chains. The configuration of a DNA strand, or other long molecule, attached to a surface at one end occurs commonly in experiments[38, 39]. The addition of a hard wall boundary excludes some conformations, which can have the effect of increasing or reducing the observed persistence length depending on how the DNA is affixed. Furthermore, the restriction imposed by the surface induces a force causing the chain to pull away. The mechanical origins of this force will be scrutinized in subsequent chapters.

Chapter three will introduce the method of phase-space sampling, using the case of a freely-jointed end-pinned polymer as an illustration. In this configuration, both average

force and fluctuations depend weakly on length. In deriving the full distribution of forces, though, we discover it is non-Gaussian and asymmetric. Large fluctuations in the force can occur more frequently than a normal distribution would predict, and the average force may not coincide with the most probable force.

Chapter four extends the method to the case of a semi-flexible chain. Building on the results of chapter three, varying persistence length allows us to describe how our derived force distributions depend on chain stiffness. We find that the profile of force fluctuations along a tethered chain falls along a universal curve, peaked at the tethered end and weakest at the free end. Exploring the effect of box confinement reveals surprising non-monotonic effects on both the average force and its fluctuation along the chain. We find that weak confinement can actually reduce both average force and fluctuation, while tight confinement increases it.

In chapter five, this methodology is applied to the isometric, or fixed-extension, ensemble. This system will be highly relevant for studying DNA looping in regulation. Furthermore, this configuration mimics DNA minicircles used to probe flexibility. We find that average force decreases with chain length, as predicted by equilibrium statistical mechanics, but that fluctuations increase. We also introduce a toy model to explain the increased force fluctuations as a function of the arbitrary coarse-graining scale. We find that even as force fluctuations grow, fluctuations in work are bounded with respect to the coarse-graining scale.

Finally, in chapter six, we summarize the results contained in this thesis. The applications and limitations of the phase-space sampling method are outlined, and some future avenues of study are discussed.

CHAPTER 2

CONFINEMENT AND EFFECTIVE PERSISTENCE LENGTH

2.1 Introduction

In this chapter, we investigate the effect of confinement on semiflexible chains in equilibrium. A polymer chain adsorbed onto a surface or bounded within a vesicle is not able to explore the full range of conformational space available to a free chain. The surrounding environment can modify the shape, size, and dynamics of polymer chains[40]. While significant work had earlier been devoted to conformations of an isolated confined polymer[41, 42, 43, 44], polymer layers acquired greater interest as properties of a single polymer proved difficult to study experimentally[45, 46, 47, 48, 49]. With the growth of single-molecule techniques, studying aspects of a polymer model with a single surface-tethered macromolecule is now realizable.

Recent experiments have studied looping in short double-stranded DNA (dsDNA), by observing surface-tethered molecules[39, 38]. These experiments used a dsDNA molecule tethered to a surface, and used single-molecule fluorescence resonance energy transfer (sm-FRET) to observe its looping behavior. This resulted in measurements of the looping probability denisty (also known as the J factor) greater than that predicted by the wormlike chain model, consistent with earlier studies[50, 51, 52]. Here, looping is taken to mean any instance of the two ends of the chain being within some specified distance of one another. The wormlike chain prediction, however, is calculated from a free chain, not one confined to a half-space as in the case of a bound dsDNA. By comparing the chain statistics between unconstrained and surface-pinned chains, we can see how they are influenced by this attachment.

For chains in the flexible limit, the effects of confinement can be explained analytically

using the reflection theorem to show that the end-to-end probability distribution shifts from a Gaussian function to its derivative[53, 54, 55]. In this instance, the mean square end-to-end displacement along the confined dimension increases by a factor of two. The size of the conformational space decreases by a factor proportional to the square root of the length of the chain. We can employ these same methods to demonstrate that a flexible chain attached at one end will loop less frequently than its free counterpart, whereas one attached to a surface at its midpoint will find its tends juxtaposed more frequently.

Such a simple treatment does not exist for chains in the semiflexible regime. In a wormlike chain model describing this type of polymer, sharper bending angles have a larger associated bending energy, and appear less often in the ensemble of conformations[56]. This implies that simple random walk statistics used for flexible chains are not applicable. Instead, Metropolis Monte Carlo methods can be employed to sample the set of chain conformations consistent with the physical constraints[57, 58]. Self-intersecting conformations are rare when the contour length is on the same order as the persistence length[59], meaning the likelihood of a given conformation depends only on its bending energy via the Boltzmann factor. This allows us to use a Gaussian sampling method[60, 61, 38] to study the set of polymer chain conformations under the influence of a confining surface.

The effect of surface confinement can be characterized in multiple ways. The impact on the looped fraction is of the most immediate relevance for the DNA experiments discussed above. However, it is also important to look at how attachment can effect the whole distribution of conformations. This can be parameterized by an effective persistence length, obtained by fitting a theoretical distribution of end-to-end distance for a free wormlike chain to the sampled values obtained from simulation. For a 186-bp dsDNA, this effective stiffness can be increased by up to 8% or decreased by over 25% depending on how the strand is attached to the surface. We can also compare the relative size of conformational space available in different pinning schemes, as a way of estimating the probability of surface contact. We find the rejection rate (the fraction of excluded conformational space) can

vary by two orders of magnitude between different pinning points, and exhibits opposite temperature dependence between middle and end-pinned schemes.

2.2 Sampling Chain Conformations

2.2.1 Notation Convention

End-to-end statistics of free chains can be characterized by a stiffness parameter $\kappa = L_P/L$, where L is the length of the chain in monomers, and L_P is the persistence length. Chains with the same stiffness are expected to have the same end-to-end distribution, independent of the number of monomers contained within each persistence length. This assumption is discussed in more detail below. The end-to-end distance, also called extension, is denoted by r, and is scaled by total length L so that it falls in the range $0 \leq r \leq 1$. The probability distribution function (PDF) of end-to-end distance is denoted as $Q(r)$, and is related to the radial distribution function (RDF) as $P(r) = 4\pi r^2 Q(r)$[62], where $P(r)$ is the RDF. Note that the radial distribution function is normalized, while the probability distribution function is not.

2.2.2 Gaussian Sampling and Boundary Conditions

The persistence length of double-stranded DNA (dsDNA) is taken to be 146 bp, which can be measured experimentally in multiple ways[19, 20, 21]. The simulated wormlike chain can be discretized into units of one monomer per base pair, for simplicity. Furthermore, the DNA chain will be assumed to be electrostatically neutral, with no self-interaction, based upon the short Debye screening length (0.8 nm) corresponding to typical physiological salt conditions ([Na^+]=150 mM). Variation of this concentration can increase or decrease the persistence length[63, 64], however this effect is only $\sim 4\%$ over the range 100 mM to 1 M [Na^+][65].

Conformations of a discrete wormlike chain are built using a rigid base-pair model[66, 38]. Three dinucleotide angles, roll (ϱ), tilt (τ), and twist (Ω), are used to define the orien-

tation of one base pair with respect to the preceding one[67]. For this study, we assume the dsDNA strand is intrinsically straight with uniform, isotropic bending fluctuations. These parameters were chosen to match the overall bending and twisting persistence lengths as molecular dynamics simulations performed with the CHARMM27 force field[68]. These correspond to 4.7° for roll and tilt angles, and 4.1° for twist angles, at 20°C. Angles were chosen using a Gaussian sampling method based upon these widths[60, 61] to generate thermally-equlibrated DNA conformations. Tangent vectors $\mathbf{d}_i$ for each monomer were obtained by rotating a unit vector by the selected angles:

$$\mathbf{d}_i = \mathbf{G}_{0,1}\mathbf{G}_{1,2}\ldots\mathbf{G}_{i-1,i}\mathbf{d}_0$$

where the matrix $\mathbf{G}$ transforms between the coordinate frames of adjacent monomers

$$\mathbf{G} = \begin{pmatrix} \cos\varrho\cos\Omega - \sin\tau\sin\varrho\sin\Omega & -\cos\tau\sin\Omega & \sin\varrho\cos\Omega + \sin\tau\cos\varrho\sin\Omega \\ \cos\varrho\sin\Omega + \sin\tau\sin\varrho\cos\Omega & \cos\tau\cos\Omega & \sin\varrho\sin\Omega - \sin\tau\cos\varrho\cos\Omega \\ -\cos\tau\sin\varrho & \sin\tau & \cos\tau\cos\varrho \end{pmatrix} \tag{2.1}$$

We used C++ on a Windows PC to perform this calculation. For free chains of one persistence length, it normally takes 600 seconds to generate 10^7 chains. This running time scales linearly with both the length of the chains and the number of chains generated.

2.2.3 Rejection Criterion

In free space, every possible conformation of a polymer chain has a full 4π steradians of solid angle accessible for global orientations. When its pinned to a surface, on the other hand, we must account for the fraction of orientations that cause the chain to intersect the bounding surface. This is done by giving the first monomer of the chain after the pinning point a random rotation, with probability distributed evenly on the surface of the unit sphere. Then the probability that a particular conformation is accepted will be equal to

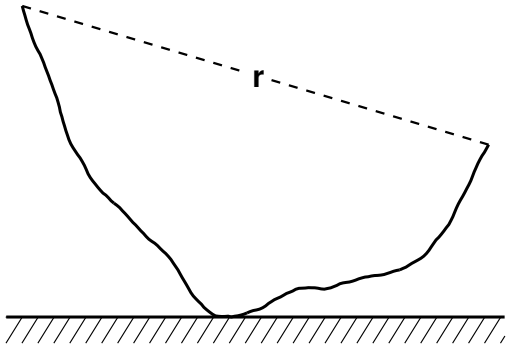

Figure 2.1: A polymer chain attached to a surface, at some point along its length. This reduces the set of accessible conformations affect the distribution of end-to-end distances r.

the fraction of global orientations compatible with the bounding surface.

The bounding surface is taken to lie in the xy-plane. Beginning with the first monomer oriented perpendicularly to this surface, along $\hat{\mathbf{z}}$, we select a random angle on the hemisphere by choosing azimuthal angle ϕ with random probability over the interval $0 \leq \phi < 2\pi$. Altitude angle θ is chosen by selecting $\cos\theta$ with uniform probability over the interval $[0, 1]$, which distributes angles uniformly over the surface of the hemisphere. This restricts us to $\theta < \pi/2$, as any larger θ would cross the boundary and be summarily rejected. This restriction is compensated by counting an additional rejection for each accepted and rejected chain, representing all of the conformations mirrored in the xy plane.

After placing the pinned link of the chain, further links are generated starting from the orientation of the pinned link and using the rotation matrices described in Equation 2.1. For a chain pinned at an internal link rather than the end, earlier links can be generated by applying the inverse of the rotation matrices. The positions of the two ends are updated in a running tally of the displacement vectors. In the event this vector crosses the boundary, the conformation can be rejected without further computation. New random angles are generated from the same distribution as before and the count of rejected chains is incremented. If no boundary crossing occurs, the chain is accepted and recorded.

Other boundary conditions, such as a spherical shell, can be implemented by calculating an aribtrary function $z(x, y)$ for the surface. This is used for comparison against the chain height when determining acceptance or rejection instead of $z = 0$. Soft surfaces can also be modeled, by introducing a potential function. In this case, chains crossing the surface are not rejected immediately, but rather with some probability according to the Metropolis criterion. Decreasing probability exponentially as potential increases results in an ensemble of states consistent with the equilibirum distribution.

2.3 Looping Rate

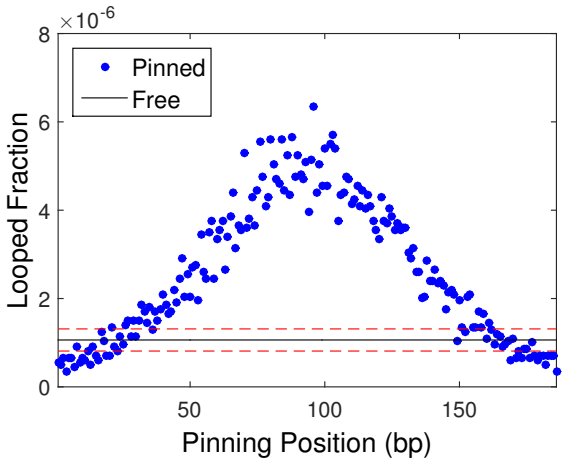

Figure 2.2: The measured looping fraction as a function of pinning position, for a 186 bp chain. When pinned near the end, the surface excludes looped conformations more often than unlooped conformations, and the looped fraction is decreased relative to the free chain. When pinned closer to the middle, unlooped conformations are rejected at a greater rate and so the looped fraction increases. For comparison, the looping fraction of a free chain is indicated by the solid horizontal line with the dashed lines representing 95% confidence intervals.

The addition of a hard boundary renders direct computation of the radial distribution function infeasible. Instead, we compute this quantity through Monte Carlo sampling of polymer conformations compatible with the surface. We employed the rigid base-pair model described above to generate conformations[38], in this case representing a 186-bp dsDNA strand, corresponding to a stiffness $\kappa = 0.78$. The looping fraction reported here

is defined as the fraction of conformations where the end-to-end distance lie within some trapping radius r_c, in this case 10 bp, or $\sim 0.05L$, where L is the total length of the strand. Mathematically, this corresponds to the end-to-end probability distribution $Q(r)$ integrated over the spherical volume of radius r_c. This is also often reported in terms of the J factor[25], equivalent to $\lim_{r\to 0} Q(r)$ in molar units. When $r_c \ll L$, the looping fraction is proportional to the J factor.

Figure 2.2 displays the effect of the surface on this quantity. For chains pinned at or near the end, the surface reduces the looping rate by approximately half relative to the free chain. A similar effect has been reported in studies attempting to describe the effect of surfaces and beads on looping in tethered particle experiments[58, 61]. Attachment near the middle, in contrast, can increase this looping rate by up to a factor of five. This skewing of chain statsitics occurs because a fully-extended chain has many orientations still available to it when attached to a surface at one end, whereas a tightly-curled conformation has comparitively few. As a result, looped conformations make up a smaller fraction of the end-pinned ensemble than the free ensemble. When attached to the surface at their midpoints, however, we find the reverse is true.

2.3.1 Temperature and Flexibility Dependence

We now investigate the dependence of this result on the flexibility of the chain. Changes in the flexibility of the chainare analogous to changing the temperature, as a chain in a heat bath at higher temperature will access larger bending angles, on average, than one at lower temperature. We restrict ourselves to a small range of variation here, as large increases in temperature will result in melting of the DNA, significantly altering its mechanical properties. Under these small changes, keeping our system within the semiflexible regime, the ratios described above do not appear to change significantly. In both middle and end-pinned schemes, decreasing the sitffness has the effect of of increasing the looping fraction by lowering the potential energy corresponding to looped states, as seen in Figure 2.3.

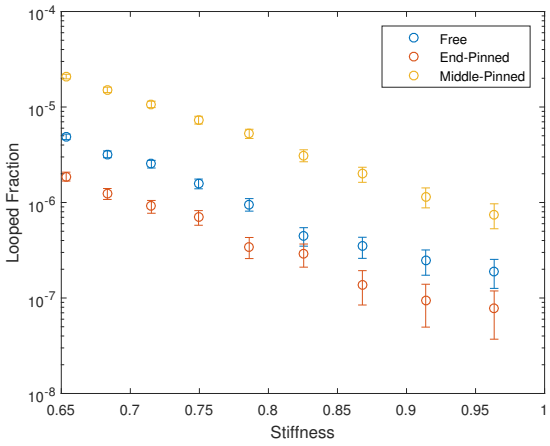

Figure 2.3: Looping fraction versus stiffness. As the chain becomes stiffer, the looping fraction decreases in all pinning schemes. Data represent 1.8×10^8 samples for the free and end-pinned geometries, and 6.0×10^7 samples for middle-pinned geometry. Bars represent 95% confidence intervals.

This is identical to the effect on free chains. The vertical distance between data points, corresponding to the ratio of looping rates, does not vary significantly in this regime.

2.3.2 Effect of Length and Scale

We can also consider changes in flexibility as changes in the length of the chain- as stiffness is defined as $\kappa = L_P/L$, increasing L has the same effect as decreasing L_P. At larger lengths, the looping rate plateaus in all geometries. Though the increased length reduces the energetic cost of bending the chain into a loop, it must eventually compete with entropic effects. For very long contour lengths, the two ends are free to explore large volumes and may be unlikely to be found near one another.

At long length scales, the advantage of the middle-pinned chain in terms of looping is expected to diminish. Over a range of one to three persistence lengths, we can already see a decrease in the ratio of the middle-pinned looped fraction to the free chain value from a factor of five to a factor of three. In the limit of a Gaussian chain, we can employ the reflection principle[53] to obtain the end-to-end distance distributions for pinned and free cases. We find the looping fraction for the middle pinned chain is expected to approach a

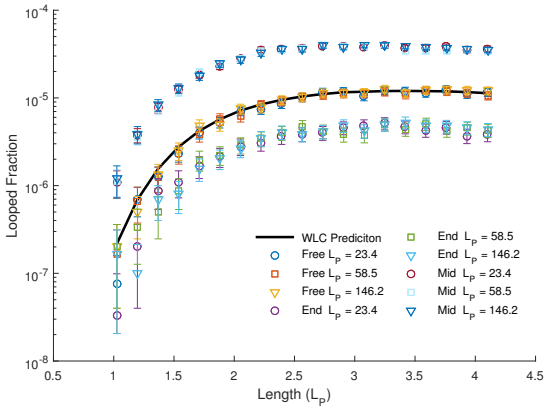

Figure 2.4: Looping rate as a function of length, in different pinning schemes. The simulations were run at three different levels of discretization of the wormlike chain, producing identical results.

factor of $\pi/2$ greater than that of the free chain. The relative ratio of looped free chains to end-pinned chains does not approach a fixed ratio, as the looped fraction falls off more quickly with the number of statistical segments N in the end-pinned geometry than in either free or middle-pinned geometries by a factor of $N^{1/2}$. As a result, the looping fraction approaches zero in the limit of long end-pinned chains. These results indicate that the J factor derived for a free chain[24, 69, 70] must be corrected to account for the geometry of surface-tethered polymers[38, 39].

We also took this opportunity to study the effect of the level of discretization. This study was initially motivated by questions about looping in single-molecule DNA experiments. For the purpose of modeling this system, it is perhaps natural to choose a single basepair as the unit of our discretized wormlike chain. However, this decision is arbitrary and so we demonstrate the independence of the looping rate results on the scale of the coarse-grained model. As the looping rate was measured at different lengths, the number of monomers per persistence length was also varied. We see in Figure 2.4 that in cases of free, middle-pinned, and end-pinned chains that the looping rate is unaffected.

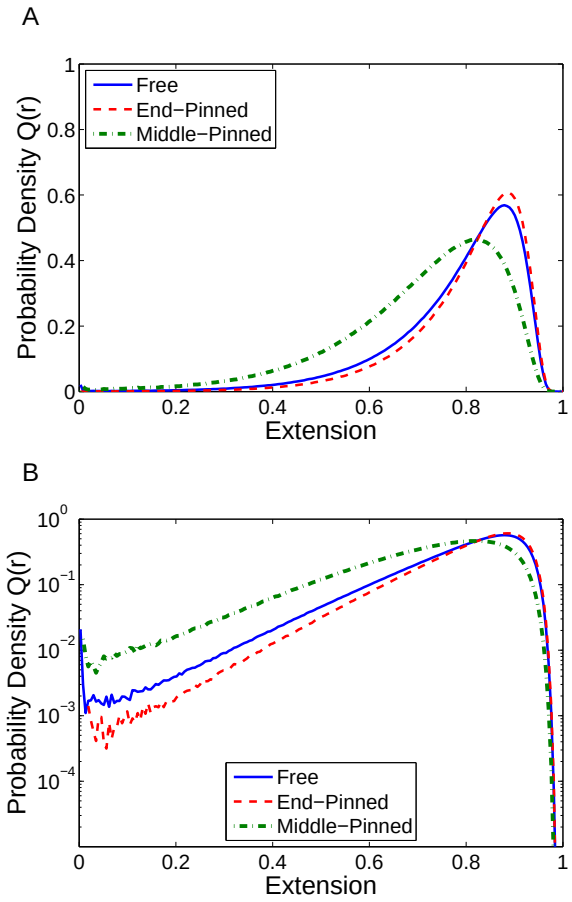

Figure 2.5: (A) Probability density of end-to-end distances for 186 bp chains in different pinning schemes. Persistence length is 146 bp. Relative to the free case (solid line), the end-pinned chain (dashed line) favors longer extensions, while the middle-pinned chain (dot-dashed line) favors shorter extensions. (B) The same data on a log scale. This better displays the significant change in probability density at small extensions for end- and middle-pinned chains.

2.4 Surface Confinement and Effective Stiffness

Apart from the change in looping rate, the effect of confinement by the surface is apparent across the entire range of extensions in Figure 2.5. Overall, the surface pushes us towards longer extensions in the end-pinned case, and toward shorter extensions in the middle-pinned case. To describe this skewing of probabilites, we attempt to fit this distribution with that of an unconstrained wormlike chain of arbitrary stiffness. The radial density

function for the free case has been studied extensively, leading to different approximations tailored for short or full extension, and for stiff and flexible limits[71, 59, 72, 73, 69]. However, each of these may break down outside of a particular regime[62]. As we seek a stiffness parameter to best represent the distribtuion over the entire extension range, we employ an expression for an exact solution as the limit of a series expansion.

2.4.1 Calculation of Q(r) for an unconstrained semiflexible chain

The Fourier transform of the end-to-end distance probability density for the free chain can be found by modelling the tangent vector at each link as a particle diffusing on the unit sphere.[74] The resulting path integral can be solved in Fourier-Laplace space to give a solution that can be conveniently expressed as an infinite continued fraction according to Mehraeen et al.[75]. The Laplace transform of the Fourier transform $\hat{Q}$ is given by

$$(\mathcal{L}\hat{Q})(p,K) = \frac{1}{P_0 + \frac{(a_1K)^2}{P_1 + \frac{(a_2K)^2}{P_2 + \ldots}}} \tag{2.2}$$

where $P_{n=0,1,2\ldots}$ and $a_{n=1,2,3\ldots}$ are defined as $P_n = p + n(n+1)$ and $a_n = n/\sqrt{4n^2-1}$. K is the reduced Fourier variable conjugate to the end-to-end distance (r). p is the Laplace conjugate of the number of statistical segments (N). The partial summation $P_n + \frac{(a_{n+1}K)^2}{P_{n+1}+\ldots}$ is denoted by j_n.

To efficiently invert both transforms, we followed the procedure given by Mehraeen[75]. The Laplace transform is inverted by identifying the poles ε_l of the expression in Equation 2.2 and employing the residue theorem.

$$\hat{Q}(K) = \sum_{l=0}^{\infty} \frac{\exp(\varepsilon_l N)}{\partial_p j_0(K, \varepsilon_l)} \tag{2.3}$$

The poles ε_l correspond to the eigenvalues of the symmetric tridiagonal matrix $-\mathbf{J}^{(0)}$, specified by diagonal entries $\mathbf{J}^{(0)}_{n,n} = n(n+1)$ and off-diagonal entries $\mathbf{J}^{(0)}_{n+1,n} = -ia_nK$,

with a_n defined as above. The infinite matrix is truncated at some size n_{cutoff}, which will give eigenvalues ε_n up to $n = n_{\mathrm{cutoff}}/4$ with sufficient accuracy.

The derivative of j_0 can be computed using a recursive relation

$$\partial_p j_n^{(+)} = 1 - \frac{(a_{n+1}K)^2}{j_{n+1}^{(+)2}} \partial_p j_{n+1}^{(+)} \tag{2.4}$$

beginning with some $\partial_p j_{n_{\mathrm{cutoff}}} = 1$ and $j_{n_{\mathrm{cutoff}}} = P_{n_{\mathrm{cutoff}}}$. After evaluating the above over a range of wavenumbers K, it is converted into a density in real space by inverting the Fourier transform

$$Q(r) = \frac{1}{(2L_p)^3} \frac{1}{2\pi^2} \int_0^\infty K^2 \frac{\sin(NrK)}{NrK} \hat{Q}(K) dK \tag{2.5}$$

using numerical quadrature.

The probability distribution of extensions obtained in this fashion can be compared to the distribution observed by counting the number of sampled chains at different extensions. Beginning with a upper and lower bound on the number of persistence lengths, a best-fit value of the chain length can be found using a golden section search[76, 77] over the parameter r. The RMS difference between the two probability distributions is used as the optimization criterion, and each successive evaluation narrows the bounds on the location of the minimum. We find the optimization function is smooth, and able to quickly guide our search procedure to a good fit value. Moreover, this value corresponds to the stiffness used to generate the end-to-end distribution, serving as a means of evaluating the correctness of the chain sampling code.

2.4.2 The Effect of Surface Pinning on Apparent Stiffness

After successfully recovering the length of free simulated chains, we then applied the same methodology to chain distributions constrained by a surface. The fit value for the intrinsic stiffness in the free chain case is $\kappa = 0.78$ for the free chain extension data. For a 186 bp

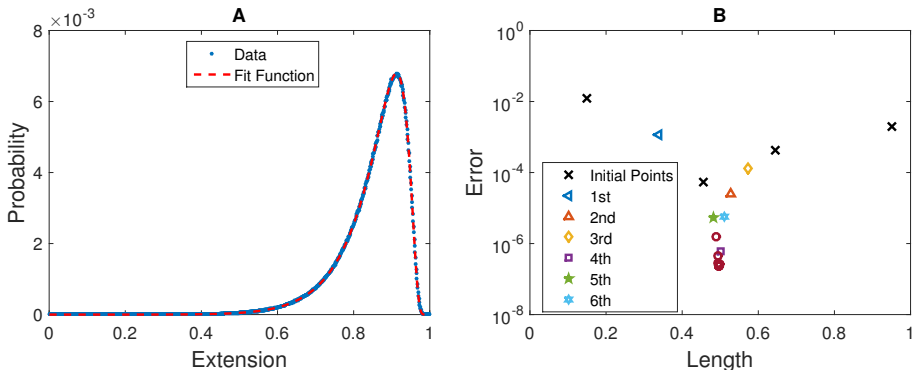

Figure 2.6: (A) Sampled probability distribution of 10^7 end-to-end distances, separated into 10^3 bins. Data represents chain of $L_P = L = 32a$. Dashed curve is a single-parameter fit obtained from end-to-end distribution function in Equation 2.5. (B) Result of golden section search. Each evaluation of residuals between end-to-end function and sample distribution narrows the bounds on location of minimum until specified tolerance is achieved.

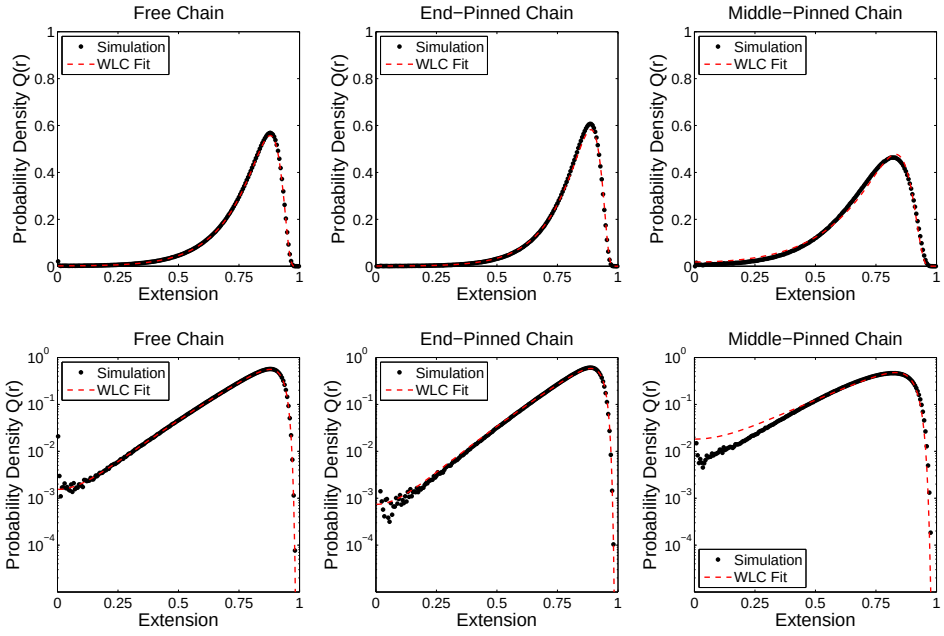

Figure 2.7: Sampled end-to-end distributions for discrete worm-like chains (black dots), alongside best fit continuous worm-like chain distributions (red dashed lines). Results represent 10^7 samples of 186 bp chains, with a persistence length of 146 bp. Linear (top row) and log (bottom row) scales are used to best display the fit at most probable and short extensions, respectively. The end-pinned chains behave like a stiffer free chain, while the middle pinned chains behave like a more flexible one. Note that the agreement between the fit and the data breaks down at short extensions for the middle-pinned chain (right panel)

chain, this corresponds to a persistence length of 146 bp, in agreement with the bending fluctuations entered into the model. For the distribution of an end-pinned chain, we observe a larger stiffness ($\kappa = 0.84$) corresponding to a persistence length of 157 bp. The middle-pinned chain displays the opposite effect, resulting in a smaller stiffness ($\kappa = 0.57$) and an apparent persistence length of 106 bp. This suggests that, considered over the whole range of extensions, the end-pinned chain is similar to a stiffer free chain, while the middle pinned chain behaves like a more flexible chain. However, the statistical behavior of the middle-pinned chain deviates from that of the more flexible free chain at short extensions.

2.4.3 Scaling of RMS Distance versus Pinning Scheme

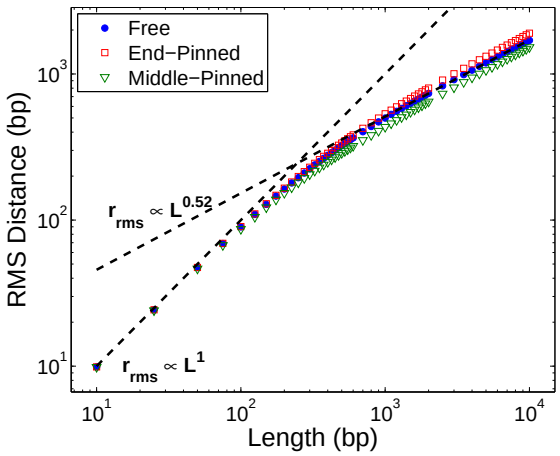

Figure 2.8: RMS end-to-end distance versus length. At short lengths, the chain does not bend significantly, and the distance scales linearly with length. At long lengths, the behavior approaches that of a random walk chain and scales as $L^{0.52}$. In the flexible limit, we would expect the distance to scale as $L^{0.5}$. The cross-over between the stiff regime and the flexible regime occurs between $1\times$ and $2\times$ the persistence length.

An alternative way of characterizing the different distributions is in terms of the root mean square (RMS) extension. We investigate the length dependence of this quantity, r_{rms}, which describes the spatial dimension of a chain similar to the radius of gyration. Figure 2.8 shows how this quantity varies with pinning scheme. In the stiff regime, r_{rms} increases linearly with the chain length in all cases, as we would expect for a stiff rod. In the semi-

flexible regime, on the order of one persistence length, the slope of this function diverges between the different pinning schemes. Beyond the stiff regime, this quantity grows at a slower rate, corresponding to the $L^{0.5}$ scaling expected for a flexible chain without self avoidance. However, the constant of proprtionality (c) differs between pinning schemes, with the end pinned chain having a longer r_{rms} rleative to the free case, and the middle pinned chain having a shorter value.

This coefficient in the scaling law $r_{rms}^2 = c \cdot (2L_pL)$ can be predicted in the flexible limit based on random walk statistics. For an unbound chain, we expect $c = 1$, while the value of 2 was found for the end-pinned chain in one dimension[42]. In three dimensions, the value $c = 4/3$ is found from the same argument[54]. For a middle-pinned chain, we calculated c to be $2 - \frac{\pi}{2}$ in one dimension, and $\frac{4}{3} - \frac{\pi}{6}$ in three dimensions. The details of this derivation can be found in Appendix A. These calculated scaling coefficients match the difference in r_{rms} exhibited by the different pinning schemes in the flexible limit.

2.5 Acceptance and Rejection Rate

To measure the size of conformational space accessible to the polymer chain, and how it changes with parameters such as length and pinning scheme, we turn to the acceptance rate. By counting the number of conformations compatible with the buonding surface, we gain perspective on the entropy difference between bound and unbound states. A chain attached in the middle can gain more entropy from detachment than an end-pinned chain[78], which would in turn imply that the middle-pinned state should be more unstable. The difference in stability could perhaps be observed by experimental measurement of the rate of dissociation from the surface under different attachment schemes.

In the flexible limit, the fraction of accepted end-pinned conformations is expected to scale with length as $L^{-0.5}$[53, 55], which is borne out by our observed scaling of $A \propto L^{-0.49}$ at long lengths. For middle-pinned chains, we instead see scaling $A \propto L^{-.96}$, approximately the square of the power law for the end-pinned case. This is reasonable, as

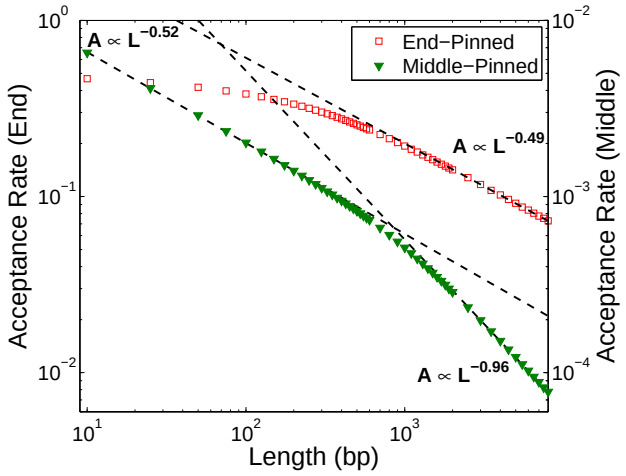

Figure 2.9: Acceptance rate (A) versus length. The end-pinned chain approaches an acceptance rate of 50% for short lengths, which we would expect for a perfectly straight rod. At long lengths, it obeys a power law $A \propto L^{-.49}$. The acceptance rate for middle-pinned chain goes as $A \propto L^{-.52}$ at short length scales, and as $A \propto L^{-.96}$. Note that the two curves are on different scales, as the acceptance rate for middle-pinned chains is significantly lower at all length scales.

there are two halves of the chain extending from the pinning position both obeying a $L^{-0.5}$ power law. At short length scales, below this regime, the acceptance rate for middle-pinned chains scales as $A \propto L^{-.52}$, while there is no clear power law for the end-pinned chains. The large gap between the two rates is principally due to the bending energy penalty associated with the internally-pinned dimer.

2.5.1 Temperature Dependence

For chains affixed to a surface in the middle or at one end, the effects of temperature variation (or flexibility variation) differ drastically. In Figure 2.10, we see that the rate of rejection increases with decreasing stiffness (defined as L_P/L) as one would expect. For a more curved conformation, fewer orientiations are compatible with the surface boundary condition relative to straighter conformations. The somewhat unexpected result is that the rejection rate decreases with increasing stiffness for a polymer atteched to the surface in the middle. While increasing the flexibility of the two free ends can lead to more conformations being rejected, increasing the bending angle at the pinning point allows the two free ends

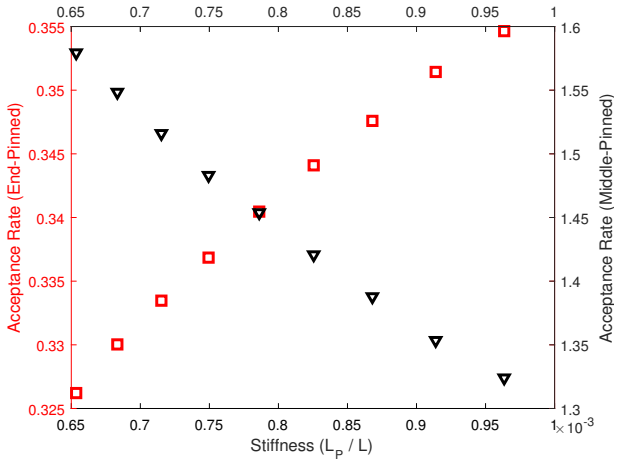

Figure 2.10: Acceptance rate (A) versus stiffness. While looping rate scales the same way in all geometries (Figure 2.3), we see two different trends for middle and end-pinned geometries.

to be directed more sharply away from the surface. Evidently this effect is more significant than the increased exclusion from the free ends. This implies that surface contact frequency should increase with temperature for an end-pinned chain, but decrease for a middle-pinned chain.

2.6 Alternative Surfaces

An impenatrable, uncurved barrier represents the simplest surface to study. However, alternative shapes may be helpful to understand DNA strands confined inside spherical vesicles. The experimental configuration for studying DNA looping referenced above may also involve a surface coated with a polymer such as polyethylene glycol (PEG), which can be more accurately modeled as via a harmonic potential representing a bed of stiff springs than with a hard wall. We briefly touch on how to model such situations and their effect on looping.

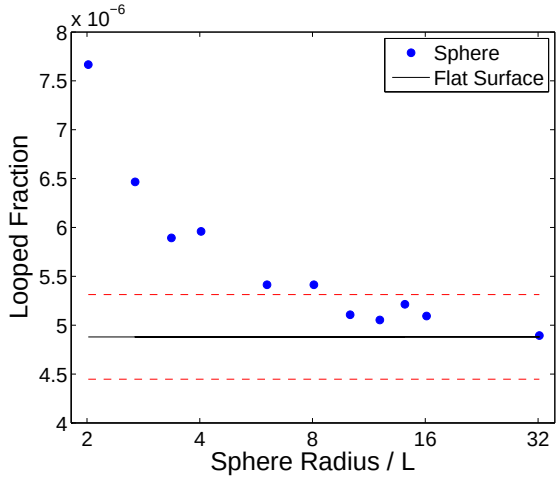

Figure 2.11: Looping rate for middle-pinned chain under spherical confinement. The radius of the sphere is expressed in units of the chain length. Small radii enhance the looping rate relative to a flat surface.

2.6.1 Spherical Confinement

Spherical confinement was studied by replacing the flat surface at $z = 0$ with a function $z = R - \sqrt{R^2 - x^2 - y^2}$, with the first link of the polymer chain remaining pinned at the origin. The parameter R is then varied over the range $2\times$ to $30\times$ the 186-bp contour length. We see in Figure 2.11 that this additional confinement increases the relative looping rate of the middle pinned chain, compared the free case. The advantage of the spherically-confined chain diminishes as the radius increases, and the configuration approaches that of a flat surface.

2.6.2 Surface Potential

Through all of the previous results, the surface has been represented by a potential stepping from 0 to ∞ across some boundary. This is computationally expedient, but it may neglect important features of a surface in a physical experiment. In practice, this surface may be charged[79] and coated with blocking agents to minimize adhesion to the surface[80]. Blocking agents such as polyethylene glycol (PEG) form a dense polymer bed that can repel molecules from the surface[81, 82]. A surface such as this would be better described

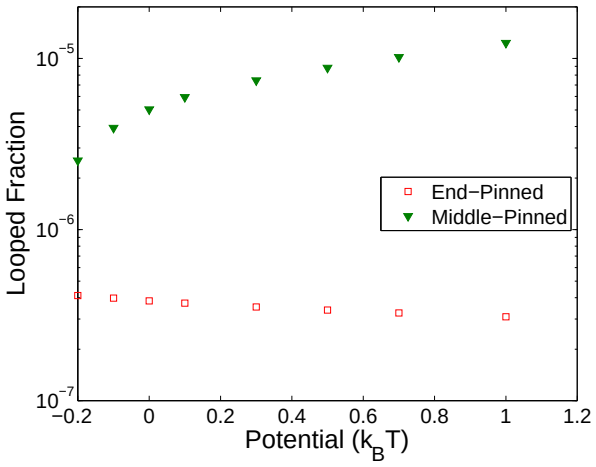

Figure 2.12: A surface potential represents blocking agents on the surface. This enhances looping in the middle-pinned case, but suppresses it in the end-pinned case.

with a soft potential function, rather than the discontinuous heaviside function employed here[83, 84, 85]. Adding an exponentially decaying repulsive surface potential, we see in Figure 2.12 that this increase the contrast between middle-pinned and end-pinned geometries, further enhancing looping in the middle-pinned case, and further supressing it in the end-pinned case.

2.7 Summary

Confinement by a surface can skew the statistics of a free wormlike chain. Depending on where the molecule is attached to the surface, this can have the effect of enhancing or suppressing looping behavior. When we compare the overall end-to-end statistics to a functional form based on a free wormlike chain, we see that pinning at the end results in a distribution resembling a stiffer free chain. In contrast, a middle-pinned chain resembles a more flexible one in some ways, though this comparison breaks down at short extensions. These effects are independent of the scale chosen to discretize the wormlike chain. We also presented scaling laws of mean square end-to-end distance as a function of length in different pinning schemes. We were able to predict the different scaling coefficients observed at large lengths using a random walk model.

We expect these results to have implications for surface-based DNA looping experiments, which have produced results that seem to contradict the wormlike behavior of dsDNA[39, 38]. Ha and Vafabakhsh used dsDNA molecules tethered internally about $1/4$ of the contour length away from the end and measured a J factor orders of magnitude higher than the wormlike chain prediciton. In contrast, Le and Kim measured a J factor from end-pinned dsDNA molecules[38] that also exceeded the wormlike chain prediction, but by no more than 10-fold. Our simulations suggest that surface confinement effects alone might account for a factor of four in the discrepancy between the two results (Figure 2.2). The overall increase in looping frequency may still require further explanation.

Surface curvature also has an affect on chain statistics[86], which may have physiological relevance. DNA attachment to membrane structures can occur inside the cell[87], and for the regulation of gene activity[88]. Furthermore, the membrane curvature is not necessarily static, but can instead by influenced by the attached polymer chains[89, 90]. Fluctuating potential functions associated with these surfaces can also have an effect on chain statistics[91, 92].

Finally, the acceptance rates reported here represent a wormlike chain affixed directly at the surface, and offer a way of measuring the partition function for the confined system relative to a free chain. By recomputing the quantities after a small change in the height from the surface, or the length of a monomer, we can approximate the change in the free energy from the change in the partition function with respect to such spatial variable. This can give us the average force, corresponding to the degree to which the confined molecule tries to break away. The next chapter focuses on this question, while also developing an alternative method for finding such forces.

CHAPTER 3

PHASE-SPACE SAMPLING

3.1 Introduction

The second law of thermodynamics predicts a system will always tend toward a state of higher entropy. When we identify different macrostates with different values of a spatial coordinate, then the tendency to move from a lower entropy macrostate to a higher entropy macrostate can be identified with an effective force along this coordinate. This is one way of deriving a macroscopic elastic force from a microscopic origin [93]. Other examples entropic force include the Casimir force [94] or depletion force [95]. The concept of entropic force has been applied to particle physics [96], and posited by some as an explanation for gravity [97, 98].

Here we use a discretized polymer chain as a model system to explore this entropic force theoretically. Employing a freely-jointed chain (FJC) model to describe a long polymer eliminates forces arising from any change in potential energy and allows us to focus on forces that are purely entropic in origin. This model effectively coarse-grains the polymer into Kuhn-length segments. Thermodynamic calculations allow us to compute the average force, from a differentiation of the free energy. The inextensibility of the chain places constraints on its motion, which the chain will naturally try to resist. This will result in a tension within the chain, analogous to a Brownian particle confined on a sphere equal to the bond length[99].

This force arises from a distribution of microstates, and as a consequence, it fluctuates as the system moves between them [100]. For a discrete wormlike chain, each of these microstates can be analyzed using classical mechanics to get a specific value for the constraint force, which allows us to translate the distribution of states into a distribution of

forces. The fluctuations in this mechanical force correspond on average to the entropic force[93, 101, 102], however their full distribution of instantaneous forces offers insight not easily obtainable from thermodynamics alone.

For an extensive thermodynamic variable, fluctuations can be calculated from the partition function when the conjugate intensive variable is fixed [103]. Examples of this include energy in a canoical ensemble or particle number in the grand canonical ensemble. The same is not true for intensive variables when te corresponding extensive variable is fixed, for instance for force in a fixed-distance ensemble [104, 105]. One can approximate this quantity by relaxing the hard constraint to a harmonic potential [106, 33], however the fluctuation will then diverge in the limit of a ery stiff potential [107]. By sampling constrained microstates, we can find an alternative way of computing this force and its fluctuation.

Further confinement of the chain, via surface attachment, volume confinement, or distace constraints between points along the chain, will represent a restriction on the states accessible to the polymer The entropic force corresponding to this reduction may manifest in the rate of detachment from a surface or escape time from confinement. This confinement-mediated force was measured for a polymer in a tube [108], and also computed using brownian dynamics simulations [109]. Heuristic arguments place the magnitude of this force on the order of $\sim k_B T/a$, where a is the length of a monomer. More involved calculations include computation of the full partition function under confinement and its derivative with respect to displacement [110]. Alternatively, this force can be calculated from the local concentration of mass points near the surface [111, 112]. For a chain tethered to a hard wall, the force increases with chain length, but saturates at $\sim k_B T/a$. For a temperature 310 K and a monomer length of 3.1 Å, this corresponds to 12.6 pN. This is consistent with 13.3 pN value for the pressure integrated over the entire surface found by Bickel *et al.*[113], where a shorter monomer length of 3 Åis assumed.

Knowing the underlying distribution of forces is important, as force fluctuations, not simply mean values, may play an important role in chemical and biological processes [15].

Rare but large deviations from the average behavior have been shown to be significant in several biological examples [114, 115]. In this chapter, we compute the distribution of forces from a flexible polymer, pinned either in space or to a surface. This can be obtained by exploring the phase-space of generalized position coordinates and their conjugate momenta. In the absence of any potential energy, it is the motion of the chain in space that will generate the fluctuating tension along the polymer. This approach represents a departure from traditional Markov chain Monte Carlo methods that only apply moves in the position space.

3.2 Partition Function and Average Force

All the information about the states accessible to a system in thermal contact with a heat bath is contained in the canonical partition function. This function, computed as an integral over all states, represents the sum of probabilities for all possible microstates of the system.

Throughout this chapter, the freely-jointed chain will be treated as flexible, but inextensible. A chain of N monomers is modeled as N point masses, corresponding to $3N$ coordinates in cartesian space. Each bond reduces the number of free coordinates by one, leaving $2N$ spatial coordinates. These can correspond to two angles for the orientation of each monomer. Note that this represents a simplification relative to the previous chapter, where tilt, roll, and twist angles were defined between each link in the chain. Representing this as a lower dimensional system via the free angular coordinates is a valid approximation when the interactions occur on a time scale which is long relative to vibrational frequencies of the bonds. We should be aware that when subjected to a sudden impact, this may break down[116].

We begin by writing down a general expression for the partition function

$$Z = \int\int \exp(-\beta \mathcal{H}(q^i, p_j)) d^{2N}q d^{2N}p \tag{3.1}$$

All momentum coordinates may range from $-\infty$ to ∞, while θ and ϕ coordinates (Figure 3.1C) for each bond range over 0 to π and 0 to 2π, respectively.

3.2.1 Constrained versus Unconstrained Ensembles

These constraints can introduce subtle differences between a representation of the system as one rigidly fixed onto a $2N$ dimensional manifold, versus a full $3N$ dimensional system kept near this manifold by a stiff potential function. Upon integrating over the momentum coordinates, we introduce a factor proportional to the square root of the metric determinant (Equation 1.13) which is absent from a treatment of the system with stiff springs in place of the rigid bonds.

This will affect the probability of visiting different spatial conformations when the system is in thermal equilibirum[117], and as a consequence the thermodynamic mean force obtained from a derivative of the free energy may differ from the average value of the mechanical constraint force. However, this discrepancy will not manifest for certain kinds of constraints, such as ones on the distance between point masses[118] used for our inextensible chain.

3.2.2 Computing Average Force

If we parameterize a partition function by a spatial coordinate ξ, then the variation of this function between surfaces of constant ξ will indicate which values of the coordinate are entropically favorable. More explicitly, we can find the entropic force from a thermodynamic calculation

$$F = \frac{\partial A}{\partial \xi} = -k_B T \frac{\partial \log Z}{\partial \xi} \tag{3.2}$$

where A is the free energy, Z is the partition function, and ξ is the generalized coordinate of interest. In this case, the reaction coordinate ξ corresponds to the length a_1 of the first monomer in the chain. We will evaluate the partition function and its derivatives at the point $a_1 = a$, where a is the length of every other monomer in the chain. In free space, the

overall partition function will be proportional to a_1^2 [99], and so the average force is given as

$$F = \frac{2k_BT}{a} \tag{3.3}$$

In the case of a chain confined to the half-space by a surface, this calculation becomes more complicated. As shown in Figure 3.1(b), lengthening the first bond of the chain increases the available conformational space in excess of the a^2 factor. By counting the fraction of chain conformations compatible with the boundary condition [119], and calculating the small increase (ΔZ) upon varying a_1 by Δa_1, we can numerically approximate Equation 3.2 in this case.

$$F \approx \frac{Z(a_1 + \Delta a_1) - Z(a_1 - \Delta a_1)}{2Z(a_1)\Delta a_1} \tag{3.4}$$

This offers a simple way to measure mean force, which can be compared against the Lagrangian constraint force averaged over the ensemble.

3.3 Sampling Forces

The preceding thermodynamic calculation can give the average force, but the distribution of forces is not easily accessible. Instead, we probe the force distribution directly by sampling possible values of the position and momentum coordinates describing the system, weighted by the appropriate probability. For each of these we can then compute the corresponding force. A large sampling of such forces will give an accurate representation of the overall average value.

3.3.1 Phase-Space Sampling

Our $2N$ spatial coordinates for the orientation of every bond will each be accompanied by a conjugate momentum coordinate. The resulting $4N$-dimensional phase space of $N-$point chains (corresponding to N monomers) can be explored with a random walk. Beginning from an arbitrary initial conformation and momentum (q^i, p_j), a random step is chosen

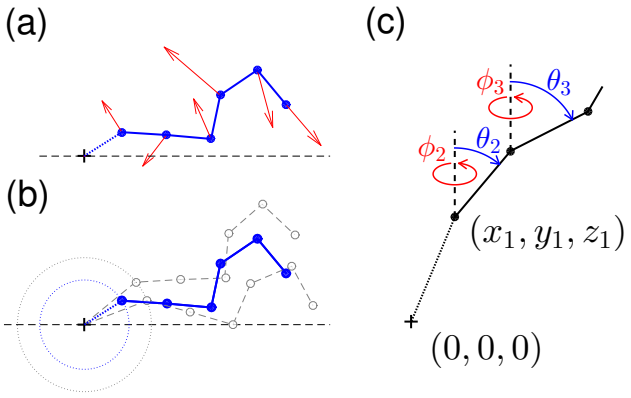

Figure 3.1: Bond force in a chain. A freely jointed chain (blue balls and sticks) is terminally pinned to a point (crosshair) and can move in full space or half space (dashed horizontal line). (a) Constraint (centrifugal) force. When freely jointed mass points of the chain move with some velocities (red arrows), the pin point will experience a force. The mass points can also exert force on the pin point when they recoil from the wall. (b) Thermodynamic (entropic) force. If the chain is pinned through a longer tether, more conformations become available. For example, the first mass point can explore a larger surface (compare blue and gray circles). Some conformations prohibited by the wall boundary can become acceptable as well when the tether is lengthened. This increase in conformational space leads to an increase in entropy, and thus a net force away from the surface. (c) Generalized coordinates in the global frame. The first three mass points of the chain are shown as black balls. The first three Cartesian coordinates (x_1, y_1, z_1) define the position of the first mass point with respect to the origin. The length of the first bond is r_1. The positions of the rest are defined by global zenith (θ) and azimuthal (ϕ) angles.

along one coordinate in phase space. We can refer to this new state as (q'^i, p'_j), though in practice only one of q^i or p_j will be altered at a time. To capture the effect of the thermal environment, this move is then accepted or rejected according to the Metropolis criterion.

$$P_{\text{accept}} = \begin{cases} 1 & \mathcal{H}(q^i, p_j) \geq \mathcal{H}(q'^i, p'_j) \\ \exp(-\Delta\mathcal{H}) & \mathcal{H}(q^i, p_j) < \mathcal{H}(q'^i, p'_j) \end{cases} \tag{3.5}$$

where $\Delta\mathcal{H}$ is the difference $\mathcal{H}(q'^i, p'_j) - \mathcal{H}(q^i, p_j)$. After an initial number of steps to equilibrate the conformation, snapshots of the system at regular intervals will produce an evenly-sampled representation of the coarse-grained chain in a heat bath.

In the case of the freely-jointed chain, the Hamiltonian will have no potential energy, and can be written as simply

$$\mathcal{H} = \frac{1}{2} p_i M^{ij}(q) p_j \tag{3.6}$$

Moves in momentum space will change the kinetic energy, but moves in position space may alter it as well. The depedence of the energy on position is captured by the mass metric $\mathbf{M}$, a bilinear form which describes how much motion corresponds to each pair of coordinates at a particular point in space. This can be found from the Jacobian

$$(\mathbf{M})_{ij} = m(\mathbf{J}^{\mathsf{T}}\mathbf{J})_{ij} = m\frac{\partial x_k}{\partial q_i}\frac{\partial x_k}{\partial q_j} \tag{3.7}$$

We find our choice of θ and ϕ angles of each monomer, measured in the global reference frame, are preferable to internal coordinates, as altering the orientation of one link in the global frame will only affect two columns of the Jacobian. Any alternative choice of coordinates will ultimately lead to the same force distribution. The contravariant form of this metric M^{ij} is defined at any given point in the space of generalized coordinates by the relationship $M^{ij}M_{jk} = \delta^i_k$.

3.3.2 Equilibrium Sampling

The procedure of perturbing the polymer coordinates in both position and momentum space may prove too time consuming for larger systems. Instead, we can consider an algorithm that perturbs the coordinates in position space and then invoke the equipartition principle to choose appropriate values for the generalized momenta. This requires us to weight each point in position space by a pseudo-potential equal to $k_B T \log(\sqrt{|M|}$ that changes probability of spatial coordinates by way of the dependence of the metric determinant on position. This captures the size of the momentum space at that location, gicing us the same distribution of spatial coordinates. The simplest example is the metric determinant in spherical polar coordinates, $\sqrt{|M|} = r^2 \sin\theta$, where the weighting makes values of θ near $\pi/2$ more probable, representing the greater surface area around the equator. Details for the computation of the metric determinant in this set of coordinates are discussed in section 3.3.4. This gives us a new Metropolis criterion for perturbing the position coordinates

$$P_{\text{accept}} = \begin{cases} 1 & \sqrt{|M(q)|} \leq \sqrt{|M(q')|} \\ \exp(\Delta|M(q)|) & \sqrt{|M(q)|} \leq \sqrt{|M(q')|} \end{cases} \tag{3.8}$$

After obtaining a sampling of conformations, each one is assigned values for its momentum coordinates as well. For a system with uncoupled degrees of freedom, like particles of gas in a box, equipartition is trivial. Constrained systems, such as a long polymer chain, are more complicated. Following the example of Jain et al[35], we introduce a set of modal coordinates ν_i at a point in space such that $p_i M^{ij} p_j = \nu_i \nu_i$. These modal coordinates are an ad hoc construct that generally lack a convenient definition in terms of the generalized coordinates, but they allow us to uncouple the degrees of freedom in momentum space. We invoke equipartion on the new set of coordinates, sampling them from a Gaussian distribution where

$$\langle \nu_i \nu_j \rangle = \delta_{ij} k_B T \tag{3.9}$$

These are then converted back into values of the original momentum coordinates. To do this, we find a Cholesky decomposition of the metric, which is a symmetric and positive definite matrix at each point.

$$\mathbf{M}^{-1} = \boldsymbol{\mu}^T \boldsymbol{\mu} \tag{3.10}$$

where $\mathbf{M}^{-1}$ is the matrix formed by the elements M^{ij}. The equivalence of the kinetic energy in the modal and generalized coordinates then gives us a triangular matrix equation to solve.

$$\frac{1}{2}\mathbf{p}^T \boldsymbol{\mu}^T \boldsymbol{\mu} \mathbf{p} = \frac{1}{2}\boldsymbol{\nu}^T \boldsymbol{\nu} \rightarrow \boldsymbol{\mu}\mathbf{p} = \boldsymbol{\nu} \tag{3.11}$$

3.3.3 Computing Forces

While a Hamiltonian formulation of the system, in terms of generalized position and momentum coordinates, was necessary to obtain the energy of the system for the purposes of phase-space sampling, we now turn to a Lagrangian description, in terms of positions and velocities, to find the force of constraint. The Lagrangian in the constrained space is given by

$$L = \frac{1}{2}\dot{\mathbf{q}}^{\mathsf{T}}\mathbf{M}\dot{\mathbf{q}} = \frac{1}{2}\dot{q}^i M_{ij} \dot{q}^j \tag{3.12}$$

which we can use obtain the equations of motion for the free coordinates.

$$M_{ij}\ddot{q}^j + \frac{\partial M_{ij}}{\partial q^k}\dot{q}^j \dot{q}^k = \frac{1}{2}\frac{\partial M_{jk}}{\partial q^i}\dot{q}^j \dot{q}^k \tag{3.13}$$

Solving for $\ddot{q}^j$ gives us

$$\ddot{q}^j = M^{ij}\left(-\dot{q}^l \dot{q}^k \frac{\partial M_{il}}{\partial q^k} + \frac{1}{2}\dot{q}^l \dot{q}^k \frac{\partial M_{lk}}{\partial q^i}\right) \tag{3.14}$$

We can express this more concisely as the geodesic equation using the Christoffel connection coefficient, and making use of the symmetry in the firstterm under exchange of l and

k

$$\ddot{q}^j = -\dot{q}^l \dot{q}^k \Gamma^j_{lk} \tag{3.15}$$

where

$$\Gamma^i_{jk} = \frac{1}{2} M^{-1}_{il} \left(\frac{\partial M_{jl}}{\partial q^k} + \frac{\partial M_{kl}}{\partial q^j} - \frac{\partial M_{jk}}{\partial q^l} \right) \tag{3.16}$$

The constraint force can be found using an undetermined multiplier λ in a Lagrangian in the larger ($2N + 1$-dimensional) space of the free coordinates and the length of the first bond.

$$L' = \frac{1}{2} \dot{\mathbf{q}}^{\mathsf{T}} \mathbf{M}' \dot{\mathbf{q}} + \lambda(a_1 - a) \tag{3.17}$$

The metric $\mathbf{M}'$ in the larger space can be expressed in a block decomposition as

$$\mathbf{M}' = \begin{pmatrix} A & B_j \\ B_i & M_{ij} \end{pmatrix} \tag{3.18}$$

where M_{ij} is the metric in the free coorindates, and the vector B_i and scalar A are obtained in a similar fashion to the other metric elements

$$A = \sum_k^N \frac{\partial \mathbf{r}_k}{\partial a_1} \cdot \frac{\partial \mathbf{r}_k}{\partial a_1} \tag{3.19a}$$

$$B_i = \sum_k^N \frac{\partial \mathbf{r}_k}{\partial a_1} \cdot \frac{\partial \mathbf{r}_k}{\partial q^i} \tag{3.19b}$$

This gives an equation of motion

$$B_i \ddot{q}^i + \frac{\partial B_i}{\partial q^j} \dot{q}^i \dot{q}^j = \frac{1}{2} \dot{q}^i \dot{q}^j \frac{\partial M_{ij}}{\partial a_1} + \lambda \tag{3.20}$$

Substituting in the expression for the generalized acceleration $\ddot{q}^i$ produces the constraint force

$$\lambda = \frac{1}{2}\dot{q}^i\dot{q}^j\left(\frac{\partial}{\partial a_1} - B^k\frac{\partial}{\partial q^k}\right)M_{ij} + \dot{q}^i\dot{q}^j M_{jk}\frac{B^k}{q^i} \tag{3.21}$$

where the vector B^k comes from applying the metric M^{jk} to the vector B_j.

3.3.4 Specifics For End-Pinned Polymer Model

All of the above applies generally to any constrained system with arbitrary coordinates. Several somplifications can be made in the specific case of interest here, of a long polymer chain described by a series of mass points and inextensible bonds. In the absence of a potential function, the constraint force can be rewritten as

$$\lambda = \sum_k^N \dot{q}^i\dot{q}^j\left(\frac{\partial \mathbf{r_k}}{\partial a} - B^K\frac{\partial \mathbf{r_k}}{q^k}\right)\cdot\frac{\partial^2 \mathbf{r_k}}{\partial q^i \partial q^j} \tag{3.22}$$

The quantity in parentheses is the derivative of the cartesian coordiantes with respect to the constrained coordinate, projected onto the space orthogonal to the constrained manifold. This is necessary, instead of simply taking the partial derivative with respect to the constrained coordinate, as the lines of constant coordinates may not be oriented orthogonal to the manifold. In such an event, our result would depend upon the (arbitrary) coordinates we used to describe the sytem within the manifold.

The manifold of constraints, orthogonal to the manifold of free coordinates, possesses a metric H_{ij}, given by

$$H_{ii} = \begin{cases} 1 & i = 1 \\ 2 & i > 1 \end{cases} \tag{3.23a}$$

$$H_{i,i+1} = -\cos(\theta_i - \theta_{i+1}) \tag{3.23b}$$

The change in H_{11} relative to the expression given b Fixman owes to the fact that the

first point of our chain is fixed. The determinant of the metric in this complimentary space matches that of the free space, but its tridiagonal strucutre makes it much easier to calculate. A recursive formula gives

$$h_k = 2h_{k-1} - \cos(\theta_k - \theta_{k-1})h_{k-2} \tag{3.24}$$

with base cases $h_0 = h_1 = 1$. h_N represents the determinant for the full metric.

Ultimately we find that the constraint force in this case is given by the expression

$$\lambda = \sum_{i=1}^{N} (\dot{\theta}_i^2 + \dot{\phi}_i^2 \sin^2\theta_i) H_{i,1}^{-1} \tag{3.25}$$

where $\mathbf{H}^{-1}$ is the matrix inverse of $\mathbf{H}$. Its elements also have a convenient definition

$$H_{ij}^{-1} = \prod_{k=i}^{j-1} \cos(\theta_k - \theta_{k+1}) h_{i-1} h'_{j+1} \tag{3.26}$$

where we have assumed $j > i$. The determinant h'_k is computed in the same fashion as h_k, but beginning at the opposite end of the chain.

$$h'_k = 2h'_{k+1} - \cos(\theta_k - \theta_{k+1})h'_{k+2} \tag{3.27}$$

where we have base cases $h'_N = 2$ and $h'_{N+1} = 1$. When the recursion reaches the fixed end, we have $h_N = h'_1$.

3.4 Trimer Example

The methodology introduced so far can be illustrated with the example of a trimer confined to a two-dimensional plane. This will be comprised of three points and two inextensible bonds, however we will assume that the first point is fixed. This will leave two free coordinates (θ_1, θ_2) for the orientation of the bonds after applying the inextensibility constraint.

The positions of the second and third mass point are given in cartesian coordinates as

$$\begin{aligned} \mathbf{r}_1 &= a_1\cos(\theta_1)\hat{\mathbf{e}}_x + a_1\sin(\theta_1)\hat{\mathbf{e}}_y \\ \mathbf{r}_2 &= a_1\cos(\theta_1) + a_2\cos(\theta_2)\hat{\mathbf{e}}_x + a_1\sin(\theta_1) + a_2\sin(\theta_2)\hat{\mathbf{e}}_y \end{aligned} \tag{3.28}$$

where a_1 and a_2 are the length of the bonds connecting the mass points. Note that in two dimensions, we expect an average force of k_BT/a, in contrast to the three dimensional case.

3.4.1 Equations of Motion

Differentiating these positions with respect to the generalized coordinates gives us a Jacobian matrix,

$$J = \begin{pmatrix} -a_1\sin(\theta_1) & 0 \\ a_1\cos(\theta_1) & 0 \\ -a_1\sin(\theta_1) & -a_2\sin(\theta_2) \\ a_1\cos(\theta_1) & a_2\cos(\theta_2) \end{pmatrix} \tag{3.29}$$

which in turn produces the metric tensor

$$M = J^TJ = \begin{pmatrix} 2a_1^2 & a_1a_2\cos(\theta_{12}) \\ a_1a_2\cos(\theta_{12}) & a_2^2 \end{pmatrix} \tag{3.30}$$

The system can explore the range $[0, 2\pi]$ in each of these angles, and the determinant of this tensor is $a_1^2a_2^2(2-\cos^2(\theta_{12}))$, where $\theta_{12} = \theta_1 - \theta_2$. This quantity is minimal for fully extended and fully contracted conformations, and maximal for conformations bent at a right angle. This indicates that in the absence of bending energy, right-angle bent conformations will be more probable than straight or hairpin conformations, per Equation 1.13. Kinetic energy can be expressed in terms of this covariant tensor and the generalized velocities, or

in terms of the contravariant metric and the generalized momenta.

$$KE = \frac{1}{2}m\left(2a_1^2\dot{\theta}_1^2 + 2a_1a_2\dot{\theta}_1\dot{\theta}_2\cos(\theta_{12}) + a_2^2\dot{\theta}_2^2\right) \tag{3.31a}$$

$$= \frac{a_2^2p_{\theta_1}^2 - 2a_1a_2\cos(\theta_{12})p_{\theta_1}p_{\theta_2} + 2a_1^2p_{\theta_2}^2}{2ma_1^2a_2^2(2-\cos^2(\theta_{12}))} \tag{3.31b}$$

The first of these yields a Lagrangian

$$\mathcal{L} = \frac{1}{2}m\left(2a_1^2\dot{\theta}_1^2 + 2a_1a_2\dot{\theta}_1\dot{\theta}_2\cos(\theta_{12}) + a_2^2\dot{\theta}_2^2\right) \tag{3.32}$$

which can be differentiated to yield two equations of motion

$$\begin{aligned} 2ma_1^2\ddot{\theta}_1 + ma_1a_2\cos(\theta_{12})\ddot{\theta}_2 &= -ma_1a_2\sin(\theta_{12})\dot{\theta}_2^2 \\ ma_2^2\ddot{\theta}_2 + ma_1a_2\cos(\theta_{12})\ddot{\theta}_1 &= ma_1a_2\sin(\theta_{12})\dot{\theta}_1^2 \end{aligned} \tag{3.33}$$

This will describe the motion of the system on the two-dimensional manifold to which it is confined. To find the constraint force required to keep it on this manifold, we turn to the Lagrangian in the higher-dimensional system where the length of the first bond is also a coordinate

$$\mathcal{L}' = \frac{m}{2}(2\dot{a}_1^2 + 2a_2\dot{a}_1\dot{\theta}_2\sin(\theta_{12}) + 2a_1^2\dot{\theta}_1^2 + 2a_1a_2\dot{\theta}_1\dot{\theta}_2\cos(\theta_{12}) + a_2^2\dot{\theta}_2^2) - \lambda_1(a_1 - a) \tag{3.34}$$

This gives us a new equation of motion in the coordinate a_1

$$\lambda_1 = m(\dot{\theta}_2^2a_2\cos\theta_{12} - 2\ddot{a}_1 - a_2\ddot{\theta}_2\sin\theta_{12} + 2\dot{\theta}_1^2a_1) \tag{3.35}$$

The velocities along the constrained coordinate ($\dot{a}_1$) is assumed to be zero, and all that remains is to solve for the undetermined multiplier λ. Here we substitute in the accelerations

we derived above in Equation 3.33, to obtain

$$\lambda_1 = \frac{2\dot{\theta}_1^2 + \cos\theta_{12}\dot{\theta}_2^2}{2 - \cos^2\theta_{12}} \tag{3.36}$$

giving the generalized force in the first bond.

3.4.2 Sampled Mechanical Force

After computing the partition function to normalize integrals, we can now find the average value and distribution of the constraint force. The two generalized velocities appearing in the equation for the constraint force are sampled from an equilibrium distribution under the assumptions outlined in Section 3.3.2. We find on integrating over velocity coordinates that they take on an average value

$$\langle \dot{q}^i \dot{q}^j \rangle = k_B T \frac{M^{ij}}{\sqrt{|M|}} \tag{3.37}$$

which gives an average force, as a function of position,

$$\langle \lambda \rangle_p = \frac{k_B T}{a} \frac{2 + 2\cos\theta_{12}}{(2 - \cos^2\theta_{12})^2} \tag{3.38}$$

The relationship varies as a function of bending angle θ_{12}. We see here that smaller values of θ_{12}, corresponding to straighter conformations, are correlated with larger force values. Integrating over all θ_{12} values, weighted by $\sqrt{2 - \cos^2\theta_{12}}$, we find an average force value of $k_B T/a$, consistent with our expectation from the partition function.

Generally, we will find that force distributions decay exponentially in probability, as opposed to obeying a Gaussian distribution. The intuitive explanation for this is that, while velocities are Gaussian distributed, forces depend on squared velocity terms, which are exponentially distributed like the energy.

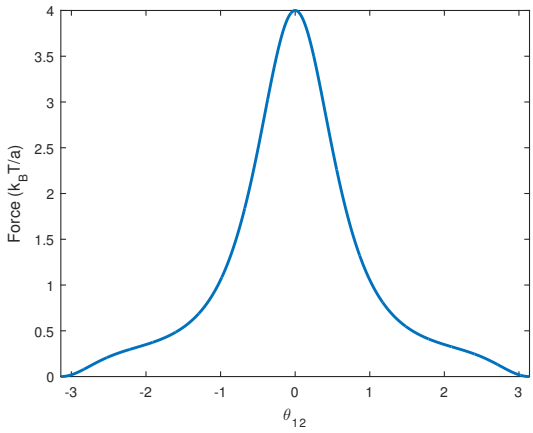

Figure 3.2: After averaging over momentum coordinates, we find an mean constraint force for every shape of the trimer. Straight conformations ($\theta_{12} = 0$ have the largest force, while those bent in a hairpin ($\theta_{12} = \pm\pi$) have zero force.

3.5 Free Chains

We now apply the same phase-space sampling method to a thermally-equilibrated flexible chain, where one end is held fixed in space. An ensemble of chain conformations is built using the equilibrium sampling method described above. For each set of position coordinates, we sample a corresponding set of momentum coordinates assuming equipartition in the modal coordinates. These two processes together give us a sampling of the phase space, which will map onto a distribution of forces.

3.5.1 Kinetic Energy Distribution

From the equipartition principle, we expect $k_B T/2$ energy on average in each degree of freedom for a system in thermal equilibrium. Due to the way the system is structured, as described by the metric tensor, this energy will not be apportioned equally between all of the mass points, as seen in Figure 3.3. Two angles in three-dimensional space means two degrees of freedom, and so we expect on average $k_B T$ energy in each mass point. We find that the first mass point has on average approximately $0.25 k_B T$ less than this, and the last point in the chain has $0.25 k_B T$ greater. The shape of this appears to be indepenedent of the

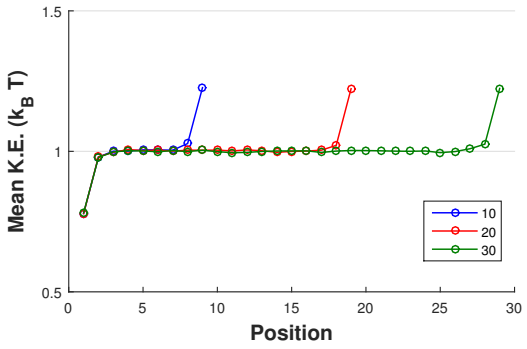

Figure 3.3: Distribution of kinetic energy for different points along the chain. The expected energy for each free coordinate is $k_BT/2$. The point closest to the pinning position has less than this, while the point at the free end has more.

length of the chain.

3.5.2 Force Distribution

The mean constraint force is toward larger a, and is the same independent of length. The value of $2k_BT/a$ is easily derived from entropic arguments[99]. Sampling the whole distribution of states allows us to calculate individual constraint forces from Lagrangian mechanics (Equation 3.21). For a single monomer, the force is always positive. Analogous to a simple pendulum, the centrigual force is proportional to the kinetic energy. The probability distribution of energy is exponentially decaying in the canonical ensemble, and so the force follows the same trend.

$$p(E)dE = e^{-\beta E}dE = \frac{a}{2}e^{-\beta a f_c/2}df_c \tag{3.39}$$

The observed distribution for longer chains, however, can take on both positive and negative values, with a sharp cusp in the probability distribution (Figure 3.4). The distribution falls off more slowly than a normal distribution about this cusp. Due to its skewness, the most probable force does not align with the mean force. The width of the distribution, measured in terms of the standard deviation or assessed visually from the probability denisty function,

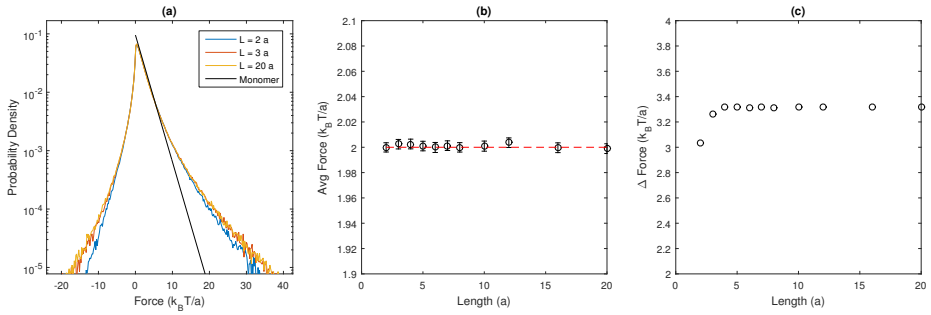

Figure 3.4: Chain pinned in full space. (a) Distribution of radial constraint forces. While the mean constraint force equals that due to a single mass point, additional links in the polymer chain increase likelihood of strong pulling forces and even introduce the possibility of compressive forces. (b) Mean constraint force and entropic force. The average force (black circles) is independent of length, and matches the predicted entropic force (red line), which results from the first mass point confined to the surface of a sphere moving to a larger radius. Results are taken from 2×10^6 samples. Bars indicate 95% confidence intervals.(c) Standard deviation of constraint forces. The standard deviation is on the same order of magnitude as the mean force, and quickly saturates as chain length increases.

exhibits little dependence on chain length beyond a small number of monomers.

As shown in Figure 3.4(a), this exponential distribution (dashed line) can explain much of the distribution of negative constraint forces. However, positive forces and a significant fraction of large negative forces deviate from this distribution. This result suggests that freely jointed mass points beyond the nearest one can have a collective impact on the pin point.

To understand combined force generation by freely jointed mass points, we examined conformations that produce anomalous force values outside the range attributable to a single monomer. As shown in Figure 3.5, instances where the force has the positive sign result from sharp bending and large velocities. For the simple case of a trimer in two dimensions, we can derive a condition for compressive force

$$\dot{\theta}_2^2 \cos(\theta_2 - \theta_1) > 2\dot{\theta}_1^2 \tag{3.40}$$

This corresponds to the case where the centrifugal force exerted by m_2 on m_1 towards the pin point is more than twice that exerted by m_1 on the pin point in the opposite direction.

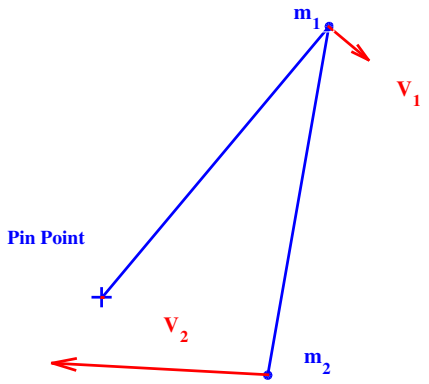

Figure 3.5: Origin of compressive forces. Some fraction of conformations (blue) and velocities (red) produce compressive forces along the first link in the chain, in contrast to the expectation from a single link acting as a pendulum. These occur when the chain is sharply bent, and the downstream mass points are moving more rapidly than those closer to the pin point ($v_2 > v_1$).

3.6 Surface Confinement

After establishing the phase-space sampling method in the case of a free chain, we turn to the example of a flexible polymer confined to the half-space by a rigid surface. Such confinemnet geometries arise commonly in biological systems [78, 120] and polymer applications [121]. We can treat this problem with the same phase-space sampling method applied above, with the added step of rejecting perturbations to the conformation that result in the chain crossing the boundary. This restricts the set of conformations accessible to the polymer, in such a way that the conformational space will increase with the length of the first bond. This is expected to induce an additional entropic force in the chain, which should be manifested in an increased average mechanical force.

3.6.1 Confined Trimer Example

As for the free chain, we begin with an illustration via the trimer. In this instance, we may find it is actually easier to treat the partition function of the three-dimensional trimer than the two-dimensional one. Defining the surface as the plane $z = 0$, we see that the zenith

angle of the first monomer ranges from 0 to $\pi/2$. The range of angles accessible to the second monomer varies with the position of the first- when the first monomer lays along the surface ($\theta_1 = \pi/2$.) then the second monomer has the range $(0, \pi/2)$ available. When the fist monomer is orthogonal to the surface, the full range $(0, \pi)$ can be accessed. More generally, the range is $(0, \pi - \theta_1)$. Applying these limits to an integral over all states gives us

$$Z = 4\pi^2 a^4 \int_0^{\pi/2} \int_0^{\pi-\theta_1} \sin\theta_1 \sin\theta_2 d\theta_2 d\theta_1 \tag{3.41}$$

The advantage of the three-dimensional chain is that all accessible z-displacements are equally probable. This allows us to rewrite the preceding integral as

$$Z = 4\pi^2 a^4 \int_0^1 \int_{-z_1}^1 dz_2 dz_1 \tag{3.42}$$

Now we consider what happens when the length of the first monomer is changed independently of the second.

$$Z(a_1) = 4\pi^2 a_1 a_2 \int_0^{a_1} \int_{-\min(a_1,a2))}^{a_2} dz_2 dz_1 = 4\pi^2 a_1 a_2 \begin{cases} a_1 a_2 + a_1^2/2 & a_1 > a_2 \\ 2a_1 a_2 - a_2^2/2 & a1 < a_2 \end{cases} \tag{3.43}$$

Differentiating this with respect to the monomer length a_1 at the point $a_1 = a_2 = a$ gives an average force

$$k_B T \frac{\partial \log Z}{\partial a_1} = \frac{7 k_B T}{3a} \tag{3.44}$$

which exceeds the value for the free chain by $k_B T/3$.

Now we turn to the phase-space sampling methods to find the distribution of mechanical forces. The constraint force is calculated as before, in Equation 3.36. The only difference is that the spatial range of integration is limited by the surface. This results in a new result for the constraint force in the constrained ensemble. After averaging over the momentum

degrees of freedom, we find the force as function of position is

$$\lambda = \frac{(2 + 2\cos\theta_{12})(3 - \cos^2\theta_{12})}{(2 - \cos^2\theta_{12})^2} \tag{3.45}$$

Carrying out the integral numerically results in a value $\langle\lambda\rangle \approx 2.20 k_B T/a$.

We find in this case that the correction described by Fixman[117, 118] is non-zero.

$$\frac{\partial A}{\partial \xi} = \langle\lambda\rangle_c - \frac{\partial \log\langle M^{-1/2}\rangle}{\partial \xi} \tag{3.46}$$

Upon applying this, we find that the average constraint force is brought closer to the force from free energy, but not into precise agreement.

3.6.2 Additional Mechanical Correction

The nature of confinement by the surface differs from a constraint on the distance between masses, as described in the introduction. The presence of a boundary limits the range of integration for spatial coordinates, resulting in an additional term in the relationship between the average constraint force and the mean force obtained from statistical mechanics, existing at the limit of the integral. This 'surface term' corresponds to the recoil force along the chain when it undergoes an elastic collision with the bounding surface. Conformations exactly touching the surface represent a set of lower dimension, and so the volume they occupy in the phase space is zero, explaining their absence from our sampling of phase space. Each impact with the surface, though, will result in an instantaneous impulse along the chain if we treat it as perfectly inextensible.

We can analyze this in detail for a dimer in two dimensions, or a simple pendulum of length a, elevated from the surface, and show that the mean effect of collisions exactly accounts for the discrepancy between the two calculations. The same reasoning is expected to apply for longer chains in three dimensions. In the case of a dimer elevated a distance h from a hard surface, our phase-space sampling will predict the same $k_B T/a$ average force

seen for a free chain in two dimensions. Differentiating the free energy derived from the partition function gives an alternate answer.

$$-k_B T \frac{d \log Z}{da} = k_B T \left(\frac{1}{a} + \frac{1}{(\pi + 2\sin^{-1}(h/a))\sqrt{a^2 - h^2}} \right) \tag{3.47}$$

If we now look at the result of an elastic collision between this dimer and the surface, we find when it impacts, the impulse along the connecting rod is given as $2ma\dot{\theta}\cot(\theta_0)$, where the angle at impact θ_0 is given by $\pm cos(-h/a)$ The frequency of collisions associated with a given trajectory will be given by $\dot{\theta}/(\pi + 2\sin^{-1}(h/a))$, and so the time-averged contribution from surface collisions will be

$$\langle I \rangle = \frac{2ma\dot{\theta}^2 \cot(\theta_0)}{\pi + 2\sin^{-1}(h/a)} = \frac{k_B T}{(\pi + 2\sin^{-1}(h/a))\sqrt{a^2 - h^2}} \tag{3.48}$$

matching the discrepancy found previously.

Computing this quantity exactly may represent an interesting mathematical exercise, however this is precisely the type of event where the assumptions underlying our rigid bond constraints break down. The results we obtain may have limited significance for accurately describing a physical system. A more thorough analysis may be necessary to garner meaningful information from the value of this discrepancy.

3.6.3 Long Confined Chains

Applying this to longer chains, we can compute the average constraint force in the presence of the surface. The surface will skew the distribution toward more fully-extended chains, which are correlated with larger force values, as we saw for the trimer. There is an additional entropic force, associated with the conformational space gained by pulling away from the surface, and we expect this to manifest in an increased average constraint force.

The average entropic force can be found from the acceptance rate of chain conformations, as used in the previous chapter. This rate will increase or decrease slightly upon

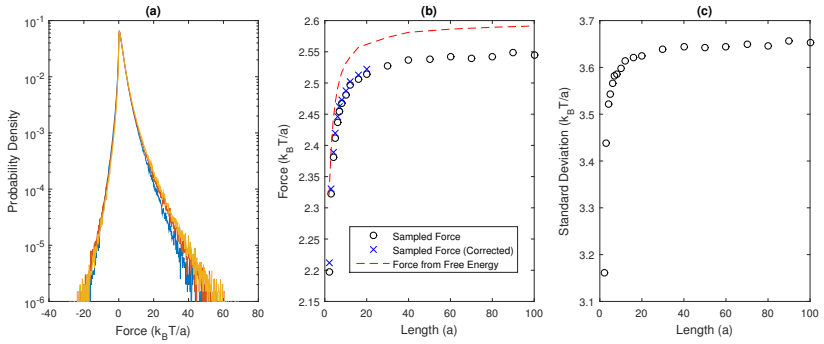

Figure 3.6: Chain pinned to plane. (a) Distribution of forces for phase-space sampled chains. The surface effect shifts the mean of the distribution slighty, but the shape is not changed relative to the free case. (b) Average sampled force before ($\circ$) and after ($\times$) applying the Fixman correction. The dashed line represents the mean thermodynamic force obtained from the partition function. (c) Standard deviation of sampled forces. As in the free chain case, the deviation is on the same order of magnitude as the mean force. It grows slightly as a function of length, quickly saturating around 5-7 links.

lengthening the first monomer of the chain. Evaluating this quantity at several values allows us to numerically approximate the derivative, and thereby obtain the entropic force. This yields a value slightly larger than the mean constraint force. Average force increases with length, but quickly saturates.

This follows the overall trend of the average force computed from sampling the rejection rate, but requires compensating for the metric determinant to bring the two into agreement. Applying the Fixman correction, we see the two results line up. An additional term, arising from collisions of the polymer with the hard surface, is necessary for exact mathematical reconcilliation of the two methods of force computation. However, for a flexible chain such as this, the effect of such collisions is small.

While the distribution of forces is shifted slightly relative to the free chain, resulting in the increased average force, the shape of the distribution does not change noticeably with the introduction of the surface. The standard deviaiton of this force (Figure 3.6 C) also follows the same trend with length as was exhibited by the free chain.

3.7 Summary

Making use of position and momentum terms in the Hamiltonian, we are able to sample the distribution of states in phase space. The equilibration of microstates in the phase space is similar to the Andersen thermostat [122] from molecular dynamics, allowing us to treat the freely-jointed chain as a canonical ensemble exchanging heat with a constant-temperature environment. A random traversal of this space using the Metropolis criterion to accept or reject moves, gives us points in phase space representing snapshots of the system as it interacts with the environment and changes its shape and momentum. Acquiring points in this space allows us to map the distribution in phase space onto a distribution of forces, for a freely jointed chain either in the full space or half space. The average of such forces is consistent with the entropic force calculated from a derivative of the partition function, differing by a known correction arising from the constrained ensemble.

In addition to describing the mechanical origin of the entropic force, this gives us access to the full distribution of forces. We find this distribution is assymetric, which results in a most probable force that does not allign with the mean force. While on average the force in the chain is tensile, rather than compressive, we find that some conformations can lead to forces that push masses closer together rather than stretching them further apart.

For a flexible chain, this distribution quickly saturates and does change shape beyond a small number (< 10) monomers. Moreover, this distribution is not Gaussian even in the long chain limit. The effect of a single degree of freedom, as in the case of a simple pendulum, is a force that falls off exponentially. As the effect of more degrees of freedom are correlated, we cannot escape this fact by considering longer chains consisting of more masses. Additionally, owing to the asymmetry of the distribution, the most probable force does not coincide with the average force. Therefore, the entropic force that we derive from statistical mechanics might not be the most relevant quantity in all circumstances.

The validity of the phase-space sampling method can be confirmed by comparison

against the thermodynamic mean force obtained from a derivative of the free energy. The two quantities line up exactly in the case of the free chain. A discrepancy emerges in the case of chains bounded by a hard surface. However, this difference is predicted theoretically and can be computed. Furthermore, we find in this case, that it has an intuitive mechanical explanation which we can demonstrate for the case of the trimer.

The same method used here can be applied to any geomtrically constrained system. In the following chapters, we will see this extended to include a bending potential at each mass point, which will influence the conformations that are chosen and contribute an additional component to the corresponding force. It can also be applied to the case of polymer chains free in solution subject to an end-to-end distance constraint.

CHAPTER 4
SEMIFLEXIBLE CHAINS

4.1 Introduction

The freely-jointed chain discussed in the previous chapter offers an interesting example to test the phase-space sampling method. The lack of a bending potential allowed us to restrict our attention to the correspondence between the mean force predicted from the increase in entropy and the average of the sampled mechanical force arising from thermal motions of the chain. By incorporating the finite persistence length of a wormlike chain, we can extend our phase-space sampling method to more physically realistic polymer chains. The elastic bending stiffness will alter the distribution of chain conformations we sample, and introduce an additional force along the length of the chain. This contribution may oppose or contribute to the force from inertial contributions alone, depending on the shape and motion of the chain.

Additionally, we can incorporate analysis of force along the chain, not just in the first monomer. The force at the anchor point has been the subject of extensive study because of its relevance for polymer detachment[123], translocation[124], and membrane curvature[113]. However, the internal force that develops along the contour, which fluctuates in thermal equlibrium[125], is also worth investigating. While we can obtain the average force at the fixed end of the chain by measuring the size of the conformational space, computing force in the same fashion for every monomer in the chain will be prohibitively time consuming. The phase-space sampling of constriant forces, on the other hand, offers us a means to accomplish this. Mapping the force distribution along the entire length of the chain may prove useful for understanding polymer stability and breakage profiles in tight confinement.

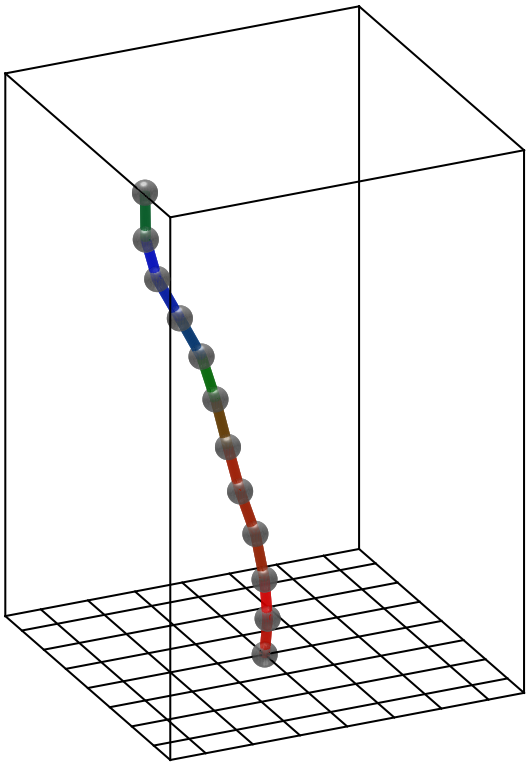

Figure 4.1: Example conformation of polymer in a box. Colors represent varying tensile force along the length of the chain.

These methods are applied first to a chain fixed at one end, but otherwise free to move about in space. The same thermodynamic arguments used to find average force in the flexible chain will apply here, independent of bond index, flexibility, or length. For a chain confined by a surface, the set of accessible conformations will be skewed toward longer extensions, as described in chapter two. This will manifest as an increase in the average sampled constraint force, but can also be observed in the first link of the chain via a finite difference derivative of the partition function. We also explore the case of additional confinement, in the form of a bounding box. (Figure 4.1) The motivation for this is to con-

sider how excluded volume interactions in a dense environment may influence the force distribution- one polymer strand in a brush will generally take on more extended counterparts than an identical strand in isolation. Like the confining surface, this is represted by flat, impenetrable walls that cause us to reject some Monte Carlo steps. This confinement is expected to have a non-monotonic effect on chain extension, and concordantly can increase or decrease the average force in the chain as it is varied.

The representation of a particular length of polymer, expressed as a multiple of the persistence length, requires choosing an arbitrary number of monomers to represent it. If the forces measured from this system are to have physical significance, they must be independent of this arbitrary choice. With this in mind, we will also look for scaling laws in both the average force and the fluctuation that will describe how they vary with the number of monomers used to represent a given persistence length.

4.2 Trimer Example Revisited

4.2.1 Constraint Forces With Potential Energy

We can return to the trimer example used to illustrate the phase-space sampling method in the previous chapter, and introduce a harmonic bending potential corresponding to a wormlike chain. Used as a Boltzmann weighting factor on the conformations, this potential will reduce the probability of sharply-bent states. The Boltzmann weighting of the straight state will be unnaffected, but as the overall partition function decreases, its likelihood will go up.

We use the same set of coordinates for the dimer example laid out in the previous chapter, and the analysis proceeds identically until we introduce a new Lagrangian with a potential function now included.

$$\mathcal{L} = \frac{1}{2}m\left(2a_1^2\dot{\theta}_1^2 + 2a_1a_2\dot{\theta}_1\dot{\theta}_2\cos(\theta_{12}) + a_2^2\dot{\theta}_2^2\right) - \frac{\alpha}{2}(\theta_1 - \theta_2)^2 \tag{4.1}$$

This can be differentiated to yield new equations of motion

$$\begin{aligned} 2ma_1^2\ddot{\theta}_1 + ma_1a_2\cos(\theta_{12})\ddot{\theta}_2 &= -ma_1a_2\sin(\theta_{12})\dot{\theta}_2^2 - \alpha\theta_{12} \\ ma_2^2\ddot{\theta}_2 + ma_1a_2\cos(\theta_{12})\ddot{\theta}_1 &= ma_1a_2\sin(\theta_{12})\dot{\theta}_1^2 + \alpha\theta_{12} \end{aligned} \tag{4.2}$$

Solving for the acceleration, we find the corresponding equations for the undetermined constraints

$$\lambda_1 = \frac{2\dot{\theta}_1^2 + \cos\theta_{12}\dot{\theta}_2^2 - \alpha\theta_{12}\sin\theta_{12}(2+\cos\theta_{12})}{2-\cos^2\theta_{12}} \tag{4.3a}$$

$$\lambda_2 = \frac{\dot{\theta}_2^2 + \cos\theta_{12}\dot{\theta}_1^2 - \alpha\theta_{12}\sin\theta_{12}(1+\cos\theta_{12})}{2-\cos^2\theta_{12}} \tag{4.3b}$$

giving generalized forces in the first and second bond. These can be decomposed into two different parts- λ_K, the portion of the constraint force arising from the kinetic motion of the mass points, and λ_P, the portion attributable to the bending potential. The form of λ_K is identical to that derived in the preceding chapter. Its distribution will differ, however, as the bending energy biases the system toward smaller values of θ_{12}. In this case, we find that the partition function Z depends on the choice of bending energy.

$$Z = \frac{2\pi a_1 a_2}{\beta m}\int_0^{2\pi}\int_0^{2\pi}\sqrt{2-\cos^2(\theta_{12})}\exp\left(-\frac{\beta\alpha}{2}(\theta_1-\theta_2)^2\right)d\theta_1 d\theta_2 \tag{4.4}$$

Carrying out the second set of integrals must be done numerically, producing the partition function dependent on bending rigidity seen in Figure 4.2. Increasing the bending rigidity decreases the probability of all the non-straight states, reducing the partition function of the system. The expectation of the force can be found as before. After integrating over momentum coordinates, we have a force as a function of bending angle θ_{12}.

$$\lambda_1 = \frac{2-2\cos\theta}{(2-\cos^2\theta)^2} - \frac{\alpha\theta\sin\theta(2+\cos\theta)}{2-\cos^2\theta} \tag{4.5a}$$

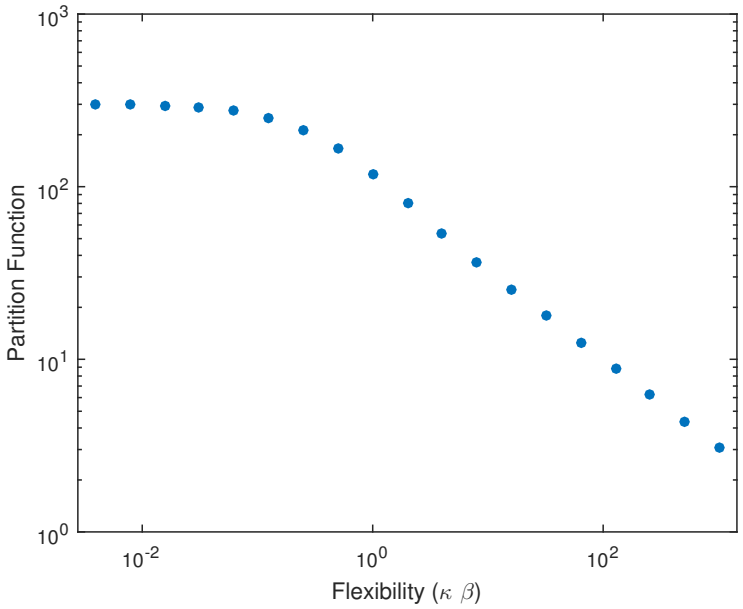

Figure 4.2: Partition function for the two-dimensional dimer. As the bending rigidity increases, the probability of bent states goes down and the partition function decreases.

$$\lambda_2 = \frac{2 - \cos\theta}{(2 - \cos^2\theta)^2} - \frac{\alpha\theta \sin\theta(1 + \cos\theta)}{2 - \cos^2\theta} \tag{4.5b}$$

Integrating this over the range $-\pi, \pi$ and weighting by the determinant factor $\sqrt{2 - \cos^2\theta_{12}}$ and the appropriate Boltzmann factor for the potential energy yields the expected average force in both bonds. Separated out into contributions arising from the potential and kinetic energy, we see that the individual components comprising λ_1 have larger magnitudes than for λ_2 (Figure 4.3 (a),(b)) Related to this, the distribution of forces is broader in the case of λ_1 as well.

These same methods can be applied to chains of arbitrary length, and under different types of confinement.

4.3 Free Chains

Our model of a polymer chain consists of $N + 1$ points of equal mass m, connected by N inextensible bonds of equal length a. Each bond is specified with a bond index (i), an

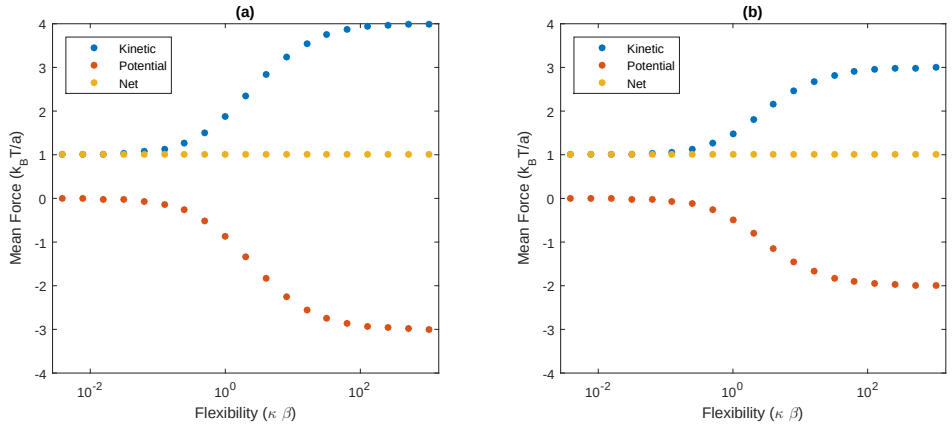

Figure 4.3: Mean force in bonds for a two-dimensional dimer. The net force is constant in all cases, however it decomposes into kinetic and potential components that increase and decrease, respectively, as rigidity increases.

integer that starts from 1 for the bond pinned to the surface and increases for each successive bond along the contour. Each mass point can be thought of as a Brownian particle confined to a sphere of radius a away from the previous point on the chain. Therefore, in the absence of any other complications, the expected force of tension in the bond is $2k_BT/a$ as shown previously [99, 126].

4.3.1 Force Components

The introduction of a bending potential means that the force values reported can arise from two different causes. Decomposing this force into contributions from the bending potential and from the inertial forces resulting from the motion of the chain, we see that the bending potential increases the average inertial force. The previous chapter demonstrated that straighter conformations were correlated with stronger inertial forces, so this result is unsurprising. The contribution from the bending potential opposes this increase, maintaining the average at a consistent value independent of the chain stiffness. The magnitude of these two balanced contributions increases with the persistence length.

Along the length of the chain, we see the magnitude of these separate components is greatest at the fixed end of the chain, and minimal at the free end of the chain. As more

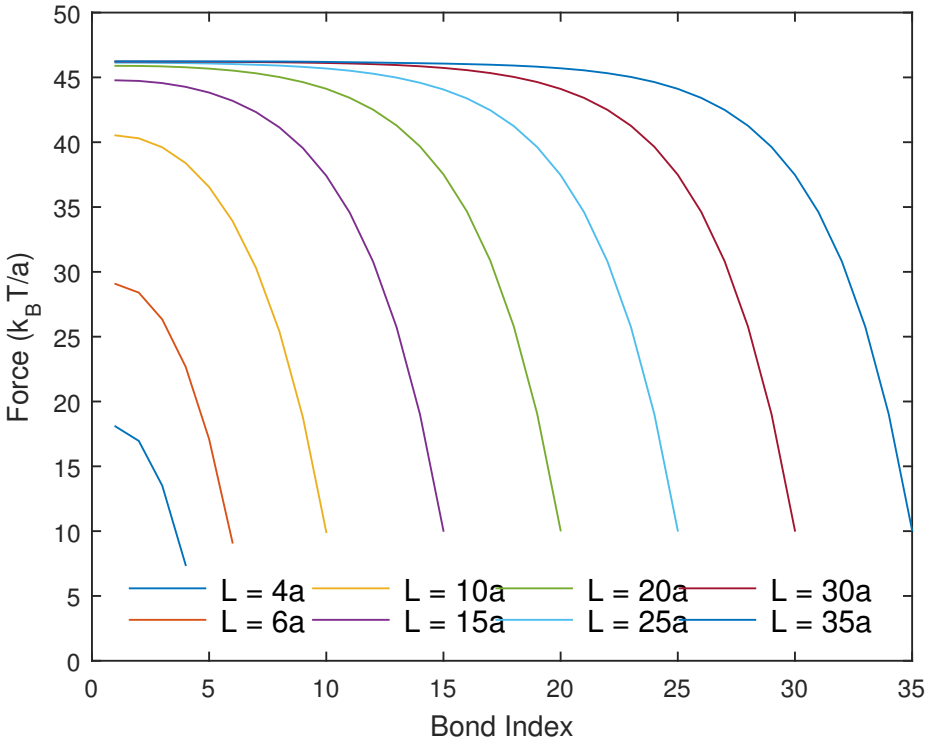

Figure 4.4: Contribution to the average force from the potential energy terms. Data taken from 10^5 samples of chains with $L_P = 23.4a$. Average kineitc contribution is in the opposite direction, shifted by $2k_BT/a$.

links are added to the chain, the magnitude of these two terms grow, though the growth stops on the order of $L = 2L_P$. Figure 4.4 shows this effect for chains with persistence length $L_P = 23.4a$. The kinetic contribution is omitted, as it is the mirror image of the potential contribution.

4.3.2 Force Distribution

At a specific bond index, the distribution of forces follows a log normal distribution. While this has the same exponential tail, favoring positive (tensile) forces as was found in the case of the flexible chain, the distribution is smoothed around the most probable value. Characterizing the width of this distribution by the standard deviation, we see that even though the average is uniform across the length of the chain, the fluctuations are largest at

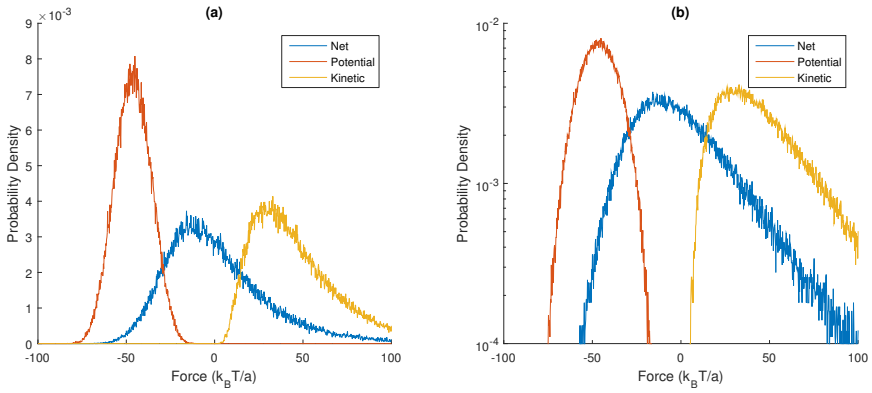

Figure 4.5: Distribution of potential and kinetic contributions to the net force in the semi-flexible chain. The distribution of the potential force falls of sharply, but the distrbution for the kinetic contribtuion has a long exponential tail favoring tensile rather than compressive forces.

the fixed end. The amplitude of fluctuations remains roughly constant through the middle of the chain, and drops of as we approach the free end. The width and height of these features remains constant above a certain length. Increasing the number of monomers only serves to increase the length of the 'plateau' region between the fixed and free end of the chain.

Varying the number of monomers per persistence length effectively allows us to alter the level of coarse graining in the representation of the system. Increasing the number of monomers per persistence length while allowing for equipartition in kinetic degrees of freedom leads to increasing energy as the represention becomes more finely grained. As a result, fluctuations in force increase.

The profile of fluctuation along the length of the polymer chain appears to follow similar curves (Figure 4.6), allowing us to describe the shape at one level of granularity as a rescaling of that found for a longer or shorter chain with a different persistence length. Increasing the persistence length of the chain causes fluctuations to increase proportional to $L_P^{0.74}$. The horizontal scale of the distribution goes as $L_P^{0.45}$.

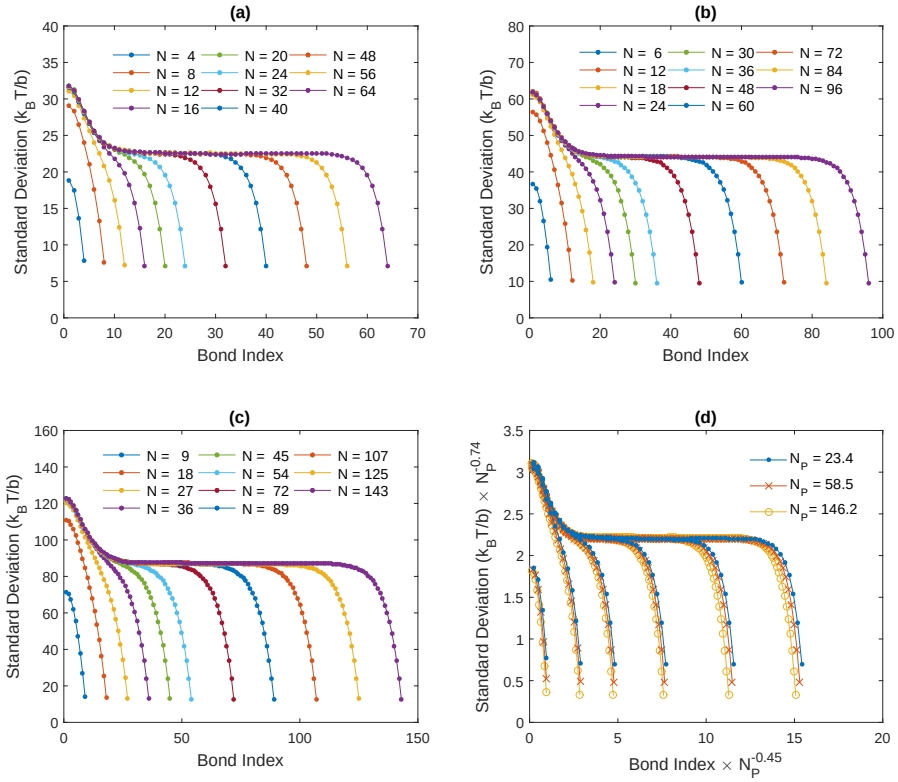

Figure 4.6: Scaling of fluctuation in free chains. The shape of fluctuations along the contour

4.4 Surface-Pinned Chains

4.4.1 Thermodynamic Mean Force

As in the case of the freely-jointed chain, an average force at the fixed end of the chain can be computed from the change in the partition function in the presence of a bounding surface. This force increases with contour length for short chains at all persistence lengths, though for longer persistence lengths we see a peak in the force at short lengths. At long contour lengths, the force decreases with the persistence length. This result can be rationalized by comparing the fractional increase in the number of conformations upon lengthening the first bond[126]. A flexible chain crosses the surface more frequently than a stiff chain, and therefore stands to gain more entropy by pulling away from the surface. Equivalently, the flexible chain creates a higher pressure field near the surface[113]. As a result, the

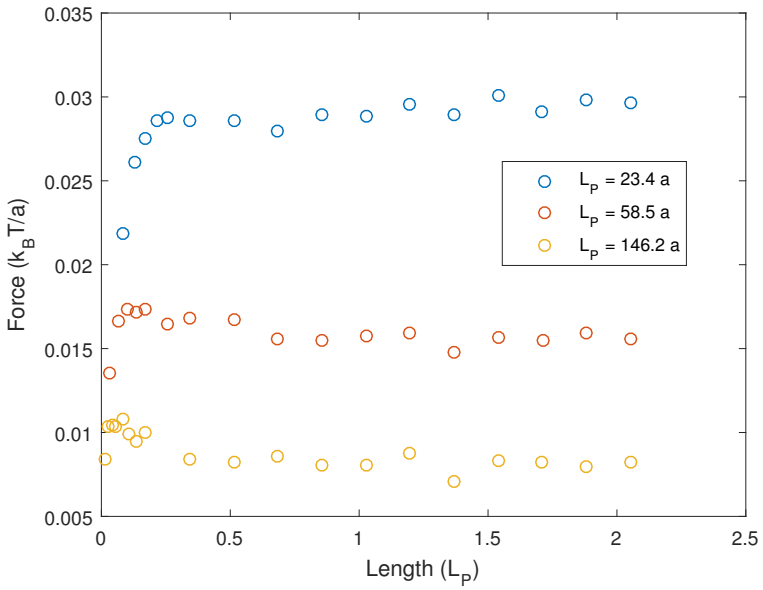

Figure 4.7: Measurement of mean force in surface-confined polymer at the base, approximated from the change in partition function. Stiff chains cross the surface less, and so exhibit a weaker force.

flexible chain exerts a stronger mean force than the stiff chain.

Intuitively, we might expect the force to scale inversely with the persistence length, as this would correspond to the same physical force after accounting for the difference in the unit of k_BT/a. We find that this is not the case, and that the force instead scales as $\approx L_P^{0.7}$. This apparent failure may be a consequence of the diverging force corresponding to the infinite potential jump at the hard boundary, or may suggest some relevant feature of the force in the chain is lost as we change the scale of our coarse-grained model.

4.4.2 Average Constraint Force

We now turn to phase-space sampling of the constraint force for comparison with our result from the partition function, and to study the distribution of forces. We find that the width and shape of the distribution is not changed noticeably from the free case, however the average is shifted towards larger tensile forces. The mean force is always the largest in the first bond, and decays to $\sim 2k_BT/a$ towards the free end, approaching the value expected

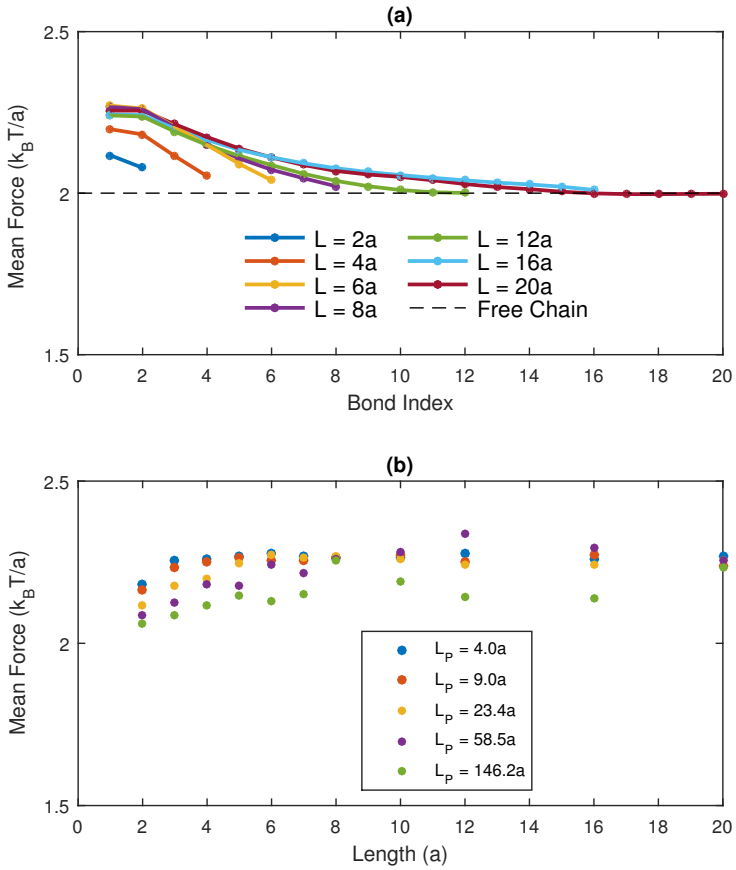

Figure 4.8: (a)Mean force as a function of position along surface-confined polymer, obtained from phase space sampling. The confined polymers exhibit a larger average force relative to a free chain, though this effect falls off further away from the pinning point. (b) Mean force in the first bond, as a function of length and flexibility. Force increases and plateaus as length increases. As persistence length increases, the effect of the surface (manifested in the degree to which the average force exceeds $2k_BT/a$) decreases.

of a chain in free space as seen in Figure 4.8(a). We can also note that the increase in force is in some cases an order of magnitude ($0.3k_BT/a$ versus $0.03k_BT/a$) larger than the effect observed on the mean force measured from the partition function.

In Figure 4.8(b), we see the mean force in the first bond increases with contour length, but only for the first few monomers. Most surprising, though, is that the average force appears to be independent of the persistence length of the simulated chain. Based upon our analysis in the preceding chapter, we hypothesize the significant disagreement between the

two measurements of average force arises from conact between the chain and the surface. As the force measured from the partition function decreases while the mean constraint force remains constant, this suggests the recoil force along the chain increases with increasing persistence length. Ths may seem at odds with the previously mentioned result of flexible chains exerting a stronger pressure field at the surface, however the relationship between the impulse imparted to the surface and the impulse along the length of the chain is complicated and may be able to reconcile both findings.

4.5 Box Confinement

After considering the case of a single confining surface, we now turn to a chain confined even further in a rectangular box. The motivation for this was to develop intuition for a polymer existing in a densly packed environment, where it will be unable to explore a wide range of conformations. Confining the simulated chains to a box of a specified width has non-trivial effects on both the mean and fluctuation of the force, as seen in Figure 4.9(a) and (b).

Loose confinement, on the order of the contour length L, decreases the average force and fluctuation (compare blue and red symbols). As we consider tighter confinement, this effect reverses, and both fluctuation and average force grow above their respective values for a surface-pinned chain without confinement. This reversal effect seems to be most pronounced for bonds in the middle of the chain. In bonds near the ends of the chain, force does not increase until confinement is on the order of the bond length (Figure 4.9(c)). The fluctuation follows a similar trend to the average force. Loose confinement decreases the width of the force distribution, while tight confinement can increase it. As a percentage of the overall fluctuation, this effect is not as significant as the effect on the average force. This dense regime is consistent with studies where tension in the first bond near the surface was shown to increase with grafting density[127].

The trend, both increasing and decreasing, correlates with the effect of confinement on

the extension of the polymer chain. Weak confinement will preferentially exclude fully-extended conformations relative to more curved ones, and as a consequence the average force will go down. As the confinemnt grows stronger, curved conformations will be ruled out and only very straight conformations will persist in the ensemble and force will increase again.

4.6 Summary

We demonstrate that, independent of the persistence length or bond index, the average force in a monomer converges on $2k_BT/a$, as can be predicted by a simple thermodynamic argument. The distribution of forces, however, can vary greatly with both of these quantities. When one end of the chain is fixed in space, the standard deviation of force is largest at this end of the chain. The fluctuation plateaus in the middle of the chain, and falls off again as we approach the free end. The shape of the decay at the fixed and free ends is independent of the length of the chain, above a certain threshold. Furthermore, the shape of this dependence appears to be universal. Rescaling the amplitude, by $L_P^{0.74}$, and the contour length by $L_P^{0.45}$, causes the fluctuations along chains of different persistence lengths to collapse onto a single curve.

Introduction of a hard wall boundary skews the distribution of conformations. This same effect occured for freely-jointed chains, as documented in the preceding chapter. More extended conformations are correlated with larger tensile forces in the polymer chain, and so the surface has the effect of increasing the average sampled constraint force. The effect of this at the base of the chain appears not to depend on the persistence length. The average force can also be obtained via a measurement of the free energy, based on the fraction of conformations compatible with the boundary condition. This produces a force that does depend on the persistence length, in contrast to the average sampled constraint force. The discrepancy between the two measurements is believed to be related to the effect of impacts on the surface, discrete events not counted in our phase-space sampling method.

Further confinement of the chain to a box produces non-monotonic effects in the force profile. Loose and tight confinement producing opposite effects on the distribution of chain extensions, with loose confinement reducing average extension and tight confinement increasing it, and the mean sampled force increases and decreases correspondingly. The effect is manifested in both the average force and the fluctuation. This result suggests that bond breakage for polymers in a tightly-packed configuration could exceed that of polymers in isolation.

Previous studies of surface-attached polymers have focused on calculating a vertical force[113, 123, 128]. The phase-space sampling approach used here allows us to instead compute the force oriented directly along the bonds- these two quantities may differ significantly, particularly under weak confinement, where the first bond may point in any direction in the half-space.

The constraint force acting in each bond might be relevant to bond stability. Our result suggests that chemically identical bonds in a polymer chain can experience a varying force spectrum depending on geometry and position. Bond dissociation in a polymer can be simulated by reactive molecular dynamics where covalent bonds are allowed to dissociate when they reach critical deformation[129, 130]. Reactive molecular dynamics studies show that bonds in the middle of a polymer break more frequently than the bonds near the free ends, and bonds break more frequently with increasing chain length[131, 132]. These results resemble ours if critical force, instead of critical deformation, is used as the bond breakage criterion. A new result obtained in our current study is that although the mean force decreases with bending stiffness, the frequency of large forces increases. Since a bond can feel more than the mean force[133], it will be interesting to measure how the detachment rate of a surface-attached polymer depends on its bending stiffness.

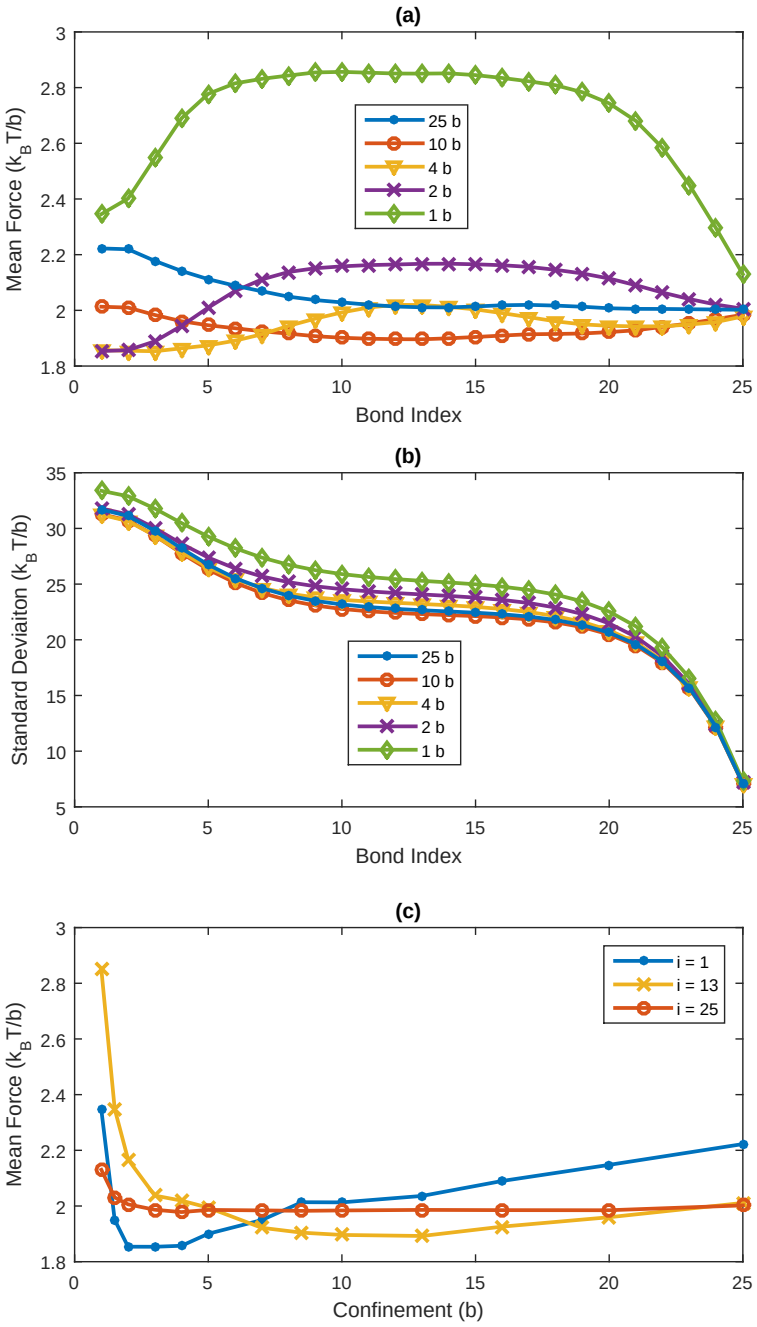

Figure 4.9: (a) Average force as a function of bond index, for different sizes of bounding boxes. Loose confinement reduces the average force, but average force increases again in tight confinement. (b) Standard deviation as a function of bond index, for different sizes of bounding boxes. Similar to the average force, loose confinement decreases fluctuation and tight confinement increases it. (c) Average force as a function of confinement at different bond index. The crossover point between increasing and decreasing force occurs at tight confinement near the pinning point, but at medium confinement near the middle of the chain Results correspond to a chain of persistence length $Lp = 23.4a$ of length $L = 25a$. Data points are averaged over 4×10^6 samples.

CHAPTER 5
THE ISOMETRIC ENSEMBLE

5.1 Introduction

Prior chapters have dealt with confinement of polymer chains by a boundary. This boundary may take different shapes, altering how the boundary skews the distribution in phase space. Here we instead consider a chain constrained by a fixed-length tether between its two ends. This represents a fundamental difference from the prior examples, in that it reduces the number of dimension for the system, instead of restricting the range the system may explore in those dimensions.

This geometry occurs frequently in nature, for instance in DNA loops[134, 135, 136], ring-like or cross-linked polymers, and cyclic macromolecules[137, 138]. While stability of such structures is generally understood in terms of free energy, it may also be impacted by local fluctuating forces acting in these systems. Therefore a full description of the distribution of forces can give us insights into mechanochemistry beyond what can be gleaned from equilibrium thermodynamics[139]. To understand this distribution, we apply the phase-space sampling methods we developed in the preceding chapters to the fixed extension, or isometric, ensemble. Generating an equilibrium distribution of polymer chains in both position and momentum space will allow us to compute the distribution of constraint forces as we vary contour length, extension, and flexibility.

We introduce a novel set of hierarchical generalized coordinates, which allow us to explore both the position and momentum space of the fixed extension ensemble. Existing Monte Carlo moves designed to preserve bond lengths, such as crank-shaft[140, 141] or fixed-end moves[142] lack well-defined partial derivatives, and therefore have no conjugate momenta. While our mechanical intuition at the macroscopic scale suggests it is the

bending elasticity that resists looping, we find that it is the thermal motion in the chain that dominates the direction of average force. The contribution of bending energy typically opposes this outward force. Both elastic and thermal contributions represent large magnitude forces, balanced so that the sum of the contributions lines up with the average force predicted from the free energy. We find that increasing contour length decreases average force, as the free energy would lead us to expect, but the additional degrees of freedom in longer chains increase the fluctuation of this force.

This trade-off between decreasing average force and increasing fluctuation is illustrated with a simple analytically solvable model. This also allows us to explore the issue of unbounded fluctuations occuring as we move to finer and finer coarse graining of our system. We find in the case of our toy model that fluctuations in work will be bounded even when fluctuations in force grow.

5.2 Coordinate System

We begin, as in the case of the terminally-pinned polymer, with an N-point chain of masses. The $3N$ cartesian coordinates of these masses are reduced to $2N$ by $N-1$ inextensibility constraints, and one constraint between the end points. Our old choice of global rotation angles as generalized coordinates is incompatible with this last constraint, and so we consider a new set of $2N$ generalized coordinates in a hierarchical fashion. At the highest level, three coordinates will describe large-scale movements of the chain. Remaining coordinates will only describe motions within one or the other half of the chain- they can be defined recursively with a set of fixed-extension coordinates being defined for each subchain as they were for the global system.

For simplicity, we assume the end-to-end vector is oriented along the z-axis. Defining the extension of the entire chain as L_1, we decompose it into two segments of length L_{10} and L_{11} (Figure 5.2). The coordinate ϕ_1 defines the azimuthal angle of the chain about the axis connecting its endpoints. The angles θ_{10} and θ_{11} are derived from the coordinates

L_1, L_{10} and L_{11}, and represent the polar angle of the segments described by L_{10} and L_{11} relative to the axis defined by L_1. These are expressed as

$$\theta_{10} = \arccos\left(\frac{L_1^2 + L_{10}^2 - L_{11}^2}{2L_1L_{10}}\right), \theta_{11} = \arccos\left(\frac{L_1^2 - L_{10}^2 + L_{11}^2}{2L_1L_{11}}\right) \tag{5.1}$$

relative to the axis of the entire chain, in a plane determined by the angle ϕ_1. This azimuthal angle ϕ_1, along with L_{10} and L_{11}, comprise the three coordinates at this level. If L_{10} and L_{11} represent single links, then those extensions are held fixed. If they represent multiple links, then they can contract or extend and the θ angles will change accordingly. At the global level, we will have five additional coordinates- three translations and two rotations of the end-to-end axis.

The advantage of these hierarchical coordinates is that the resulting metric tensor will be sparse. While the size of the tensor will scale as N^2, the number of non-zero entries will scale as $N \log N$ (Figure 5.1). This will greatly expedite computing the matrix and its derivatives.

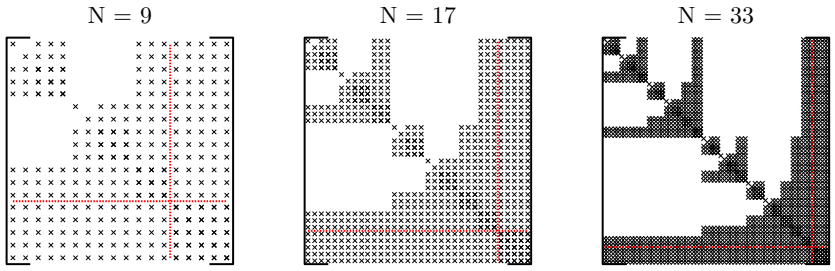

Figure 5.1: Schematic of non-zero matrix elements for different values of N. Blank spaces represent values which are always zero, $\times$'s represent values which may be non-zero. The red dashed line separates the internal and global coordinates. The pattern of non-zeros at $N = 2^l + 1$ is repeated twice within the pattern for $N = 2^{l+1} + 1$, creating a fractal structure.

5.2.1 Crankshaft Rotation Moves

A crankshaft rotation move alters one of the azimuthal angles ϕ. On a set of points, it will be defined by rotating the interior points about the axis connecting the end points.

An example is found in Figure 5.2(A). This will preserve all the interior distances, and the overall end-to-end vector for the subchain. Crankshaft moves may serve as part of a complete set of generalized coordinates, provided the intervals they span do not partially overlap. Crankshaft angles of disjoint subchains, or a subchain that is entirely contained within another, may be altered independently, but angles for partially overlapping subchains may not.

5.2.2 Expansion Moves

Expansion moves will come in two varieties. Defining the midpoint of a set of vertices $\mathbf{r}_i \ldots \mathbf{r}_{i+l}$ as $\mathbf{r}_{i+k}$ where k is halfway to l, rounded up ($k = \lceil l/2 \rceil$), one move will expand or contract the points $\mathbf{r}_i \ldots \mathbf{r}_{i+k}$ by changing the angle at their midpoint $\mathbf{r}_{i+j}$ where $j = \lceil k/2 \rceil$ while simultaneously rotating the points $\mathbf{r}_{i+k+1} \ldots \mathbf{r}_{i+l-1}$ about the point $\mathbf{r}_{i+l}$ to preserve the interior distances $|\mathbf{r}_i - \mathbf{r}_{i+l}|$ and $|\mathbf{r}_{i+l} - \mathbf{r}_{i+k}|$, as in Figure 5.2 B. The other expansion move will preserve the distances $|\mathbf{r}_i - \mathbf{r}_{i+l}|$ and $|\mathbf{r}_i - \mathbf{r}_{i+k}|$ while expanding or contracting the chain between $\mathbf{r}_{i+k}$ and $\mathbf{r}_{i+l}$, as in Figure 5.2 C. These expansion moves are similar to other algorithms based on solving the inverse kinematic problem[143, 144, 145], which work by applying a stochastic rotation step on one segment and applying a deterministic rotation step on another segment to close the chain.

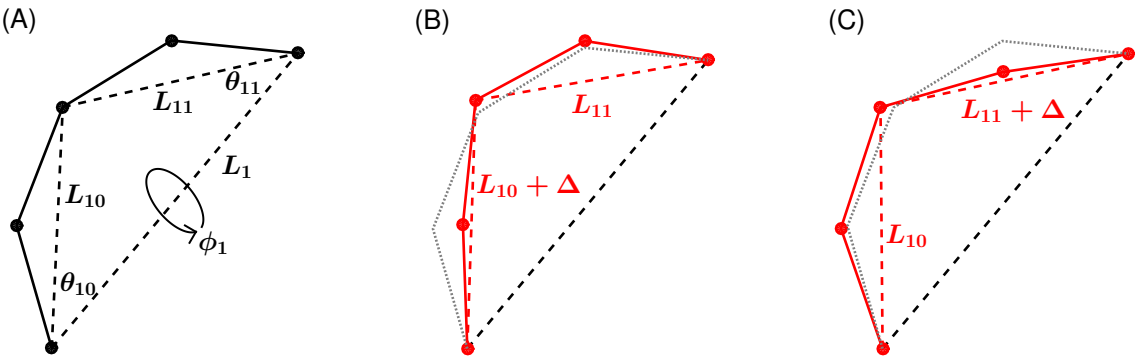

Figure 5.2: Generalized coordinates at one level. In (A), the crankshaft rotation preserves all internal end-to-end distances. In (B), the extension L_1 is preserved while L_{10} is altered. In (C), L_1 remains fixed while the extension L_{11} is changed. The same set of coordinates will exist within the subchains L_{10} and L_{11}.

We can then find the number m of generalized internal coordinates for a chain of $N-1$ links and N points, using a recursive formula

$$m(N) = \begin{cases} 0 & N < 3 \\ 1 & N = 3 \\ 2 + m(3) & N = 4 \\ 3 + m(\lfloor N/2 \rfloor + 1) + m(\lceil N/2 \rceil) & N > 4 \end{cases} \tag{5.2}$$

This can easily be shown to yield $m(N) = 2N - 5$ for all $N \geq 3$. Adding in the five global coordinates gives a full set of $2N$ generalized coordinates.

5.2.3 Cartesian and Generalized Coordinates

The relationship between the Cartesian coordinates describing the positions of point masses, and the generalized coordinates in the isometric ensemble can be described in terms of rotation matrices and translations. Rotations about all three cartesian axes can be defined as

$$\mathbf{R}_x(\alpha) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\alpha & -\sin\alpha \\ 0 & \sin\alpha & \cos\alpha \end{pmatrix} \tag{5.3a}$$

$$\mathbf{R}_y(\beta) = \begin{pmatrix} \cos\beta & 0 & \sin\beta \\ 0 & 1 & 0 \\ -\sin\beta & 0 & \cos\beta \end{pmatrix} \tag{5.3b}$$

$$\mathbf{R}_z(\gamma) = \begin{pmatrix} \cos\gamma & -\sin\gamma & 0 \\ \sin\gamma & \cos\gamma & 0 \\ 0 & 0 & 1 \end{pmatrix} \tag{5.3c}$$

The position of the first mass point of the chain will be given by three global cartesian coordinates.

$$\mathbf{r}_0 = \begin{pmatrix} c_x & c_y & c_z \end{pmatrix} \tag{5.4}$$

The last point of the chain will be defined in terms of the position of the first point, as well as an extension and two global rotations. $z-$rotations are not included, as rotation of the chain about the axis connecting the first and last point will have no effect on the position of the last mass point.

$$\mathbf{r}_{N-1} = \begin{pmatrix} c_x & c_y & c_z \end{pmatrix} + \mathbf{R}_x(\alpha)\mathbf{R}_y(\beta) \begin{pmatrix} 0 & 0 & L_1 \end{pmatrix} \tag{5.5}$$

Intervening points can be described with additional rotation matrices. Here the angle θ_{10} is determined from the extensions L_1, L_{10} and L_{11}, as described in Equation 5.1.

$$\mathbf{r}_{N/2} = \begin{pmatrix} c_x & c_y & c_z \end{pmatrix} + \mathbf{R}_x(\alpha)\mathbf{R}_y(\beta)\mathbf{R}_z(\phi_1)\mathbf{R}_y(\theta_{10}) \begin{pmatrix} 0 & 0 & L_{10} \end{pmatrix} \tag{5.6}$$

The first and last point of the chain will be denoted as $\mathbf{r}_1$ and $\mathbf{r}_N$, respectively. Rotations and translations can be used to define the midpoint of any other two points. Given two points $\mathbf{r}_i$ and $\mathbf{r}_j$, we define L_{ij} as the distance between them and $\mathbf{R}_{ij}$ as the rotation matrix relating their common reference frame to the global reference frame. We can then define the midpoint between them as

$$\mathbf{r}_k = \mathbf{r}_i + \mathbf{R}_{ij}\mathbf{R}_z(\phi_{ij})\mathbf{R}_y(\theta_{ik}) \begin{pmatrix} 0 & 0 & L_{ik} \end{pmatrix} \tag{5.7}$$

in a more general version of Equation 5.6.

5.2.4 Computing Metric Tensor Elements

The Jacobian for the system is found by differentiating the Cartesian coordinates with respect to the generalized coordinates. This can be done using the relations described above,

putting the cartesian coordinates in terms of the generalized ones. In the broadest sense, one element of the metric tensor is given by

$$M_{ij} = \sum_{m=1}^{N} \frac{\partial \mathbf{r_m}}{\partial q^i} \cdot \frac{\partial \mathbf{r_m}}{\partial q^j} \tag{5.8}$$

This can be simplified, as most positions do not depend on most generalized coordinates. Both coordinates must affect some of the same mass points to have a non-zero overlap, meaning one is an ancestor of the other, or that they exist at the same node in the hierarchy.

5.3 Monte Carlo Sampling

The coordinates described above can be used to navigate the space of fixed-extension conformations. This can be approached in the same manner as for the end-pinned chain, with a coordinate and step size chosen at random. The change in energy, including the metric correction, will be computed and the step accepted or rejected according to the Metropolis criterion.

$$P_{q\to q'} = \begin{cases} 1 & U(q') - \frac{1}{2}\log M(q') \leq U(q) - \frac{1}{2}\log M(q) \\ \exp(-\beta\Delta U)\sqrt{\frac{M'}{M}} & U(q') - \frac{1}{2}\log M(q') > U(q) - \frac{1}{2}\log M(q) \end{cases} \tag{5.9}$$

Repeating this process produces an ensemble of states corresponding to thermal equilibrium. Implementing this efficiently in a computer simulation can be done using a hierarchical, tree-based algorithm. Each trial move will have to evaluate the change in potential energy, and the change in metric determinant. This is simplified by the fact that each move in one of the hierarchical coordinates will change no more than four of the internal bending angles.

5.4 Force Computation

The constraint force required to keep the system at a fixed extension can be found from the equation of motion in the constrained coordinate q^r. Beginning with Lagrangian $\mathcal{L}$ in the $2N+1$ dimensional space of free coordinates q^i and the constrained extension, this is given by

$$\mathcal{L} = \frac{1}{2}\dot{q}^i \mathcal{M}_{ij} q^j - U(q^i, q^r) - \lambda(q^r - r_0) \tag{5.10}$$

where λ is an undetermined multiplier corresponding to the constraint force and r_0 is the constrained value of the end-to-end distance. The metric ($\mathcal{M}_{ij}$) here is in the larger, $2N+1$ dimensional space. Represented in block form, this is equivalent to

$$\mathcal{M} = \begin{bmatrix} A & \mathbf{B}^T \\ \mathbf{B} & \mathbf{M} \end{bmatrix} \tag{5.11}$$

where

$$B_i = \sum_{k=1}^{N} m_k \frac{\partial \mathbf{r}_k}{\partial q^i} \cdot \frac{\partial \mathbf{r}_k}{\partial q^\xi} \text{ and } A = \sum_{k=1}^{N} m_k \frac{\partial \mathbf{r}_k}{\partial q^\xi} \cdot \frac{\partial \mathbf{r}_k}{\partial q^\xi} \tag{5.12}$$

The metric M is defined on the free coordinates from the Jacobian, as described above.

Note that, unlike the case of bond constraints in end-pinned chains, we cannot neglect the change in bending potential with respect to our constrained coordinate here.

$$\lambda = B_j \ddot{q}^j + \dot{q}^j \dot{q}^k \frac{\partial B_j}{\partial q^k} - \frac{1}{2}\dot{q}^j \dot{q}^k \frac{\partial M_{jk}}{\partial r} + \frac{\partial U}{\partial r} \tag{5.13}$$

This expression requires that we also know the generalized accelerations of the free coordinates, which are obtained from another set of equations of motion.

$$M_{ij}\ddot{q}^j = -\dot{q}^j \dot{q}^k \frac{\partial M_{ij}}{\partial q^k} + \frac{1}{2}\dot{q}^j \dot{q}^k \frac{\partial M_{jk}}{\partial q^i} - \frac{\partial U}{\partial q^i} \tag{5.14}$$

Ultimately, this gives us an equation for the constraint force in terms of generalized posi-

tions and velocities only.In terms of the contravariant form of the metric tensor M^{ij} defined by $M_{ij}M^{jk} = \delta_i^k$, we can express the constraint force as

$$\lambda = \underbrace{-M_{ik}\dot{q}^i\dot{q}^j\frac{\partial B^k}{\partial q^j} + \frac{1}{2}\dot{q}^i\dot{q}^j\left(\frac{\partial}{\partial q^\xi} - B^k\frac{\partial}{\partial q^k}\right)M_{ij}}_{\text{Inertial}} - \underbrace{\left(\frac{\partial}{\partial q^\xi} - B^k\frac{\partial}{\partial q^k}\right)U}_{\text{Elastic}} \tag{5.15}$$

where the vector B_i is converted from covariant to contravariant using the metric M^{ij} on the unconstrained subspace. The first and second terms, which depend on the kinetic energy of the polymer chain, are categorized as the inertial force (or entropic force in an average sense), whereas the third term proportional to the bending energy with no velocity dependence is categorized as the elastic force. Force computation can also be accelerated with a hierarchical, recursive algorithm. Generally, this will be broken down into the steps of choosing appropriate generalized velocities, computing the generalized accelerations, and computing the generalized force.

As in the case of the end-pinned chain, computing constraint forces will require us to chose Gaussian-distributed random velocities in a set of modal coordinates. These are then transformed into velocities in the generalized coordinates by way of the Cholesky factorization of the metric tensor. This is done via a recursive algorithm, similar in structure to that used for find and updating generalized coordinates in the Monte Carlo step.Off-diagonal terms of the metric tensor will be computed with a branching recursion that efficiently covers the fractal landscape of non-zero entries. After obtaining a triangular factorization of the metric, modal velocities can be converted into values in the generalized coordinates by back-substitution. This same factorization will be reused to solve the equation for the constraint (Equation 5.15) once these kinetic terms and derivatives of potential energy have been computed.

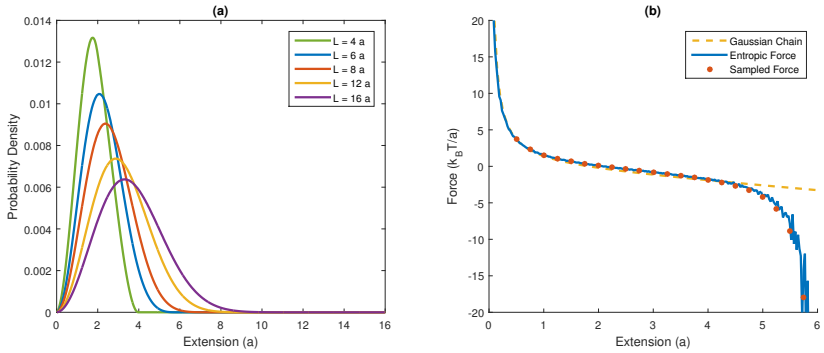

Figure 5.3: Average force for flexible chains. (a) The distribution of end-to-end distances for the flexible chain of different lengths. A geometric factor favors larger extension, while the Gaussian statistics of a random walk favor shorter extensions. (b) Average force derived from a Gaussian chain, and from a simulated inextensible chain for a chain of length $6a$. The two agree at short extensions, but diverge when the inextensible chain is taut. The average sampled force agrees with this result.

5.5 Flexible Chains

We can begin by applying this methodology to a freely-jointed chain at fixed extension. The entropic force can be estimated in this case from a Gaussian distribution, corresponding to a random walk, as described in Equation 1.4. This is an adequate description at short extensions, but breaks down near the full extension of the chain, as the Gaussian chain will continue out indefinitely with decreasing probability whereas the inextensible chain will abruptly cut off when it reaches full extension. For a more accurate non-analytical approach we can sample the distribution of extensions by generating random walks of a specifed length. The resulting end-to-end distribution (Figure 5.3(a)) can be treated as an extension-dependent partition function and differentiated to find an average force. Both of these quantities agree with the sampled mechanical force at shotrt extensions (Figure 5.3(b)) but only the random walk result lines up well at large extensions. At fixed extension, we can see the effect of increasing the length of the chain, in Figure 5.4. When the chain is only slightly larger than the extension, the chain can increase entropy by contracting. Accordingly, there is a net negative force. As the chain gets significantly longer than the extension, this entropic effect decreases. Eventually, the chain can increase entropy

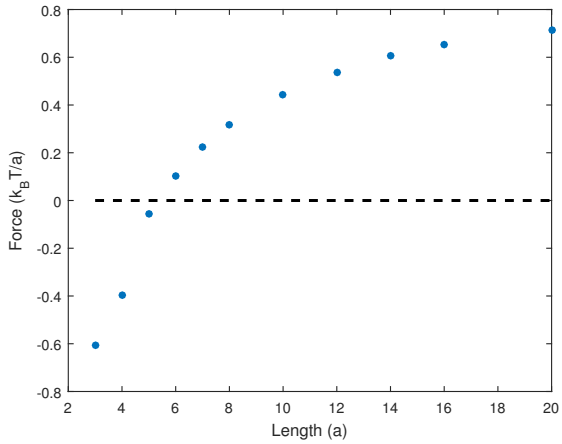

Figure 5.4: Average sampled force at fixed extension $2a$, as the contour length of the chain increases from $3a$ to $20a$. The force switches from positive to negative as the most probable extension goes from below $2a$ to above.

by increasing extension, owing to the geometric weighting by r^2, and the force becomes positive.

5.6 Constraint Force versus Force From Free Energy

For comparison and validation of the sampled instantaneous constraint force, we also compute a force from the free energy. The generalized force $\overline{f}$ corresponding to the end-to-end distance coordinate can be found from the partition function, parameterized in terms of the extension. The relative magnitude of this function is found from the end-to-end probability density function, and so we have

$$\overline{f} = \frac{\partial A(r)}{\partial r} = -k_B T \frac{\partial \log P(r)}{\partial r}, \tag{5.16}$$

In Figure 5.5B, we see this result (solid lines) overlaid with the average of the sampled constraint force $\langle \lambda \rangle$ (open circles). The generalized force was calculated using a semi-analytical expression for $P(r)$ derived by Mehraeen et al. [75], described in Chapter Two. Both methods agree across a range of extensions, evaluated for different contour lengths. This shows our phase-space sampling method is consistent with predicitions from equilib-

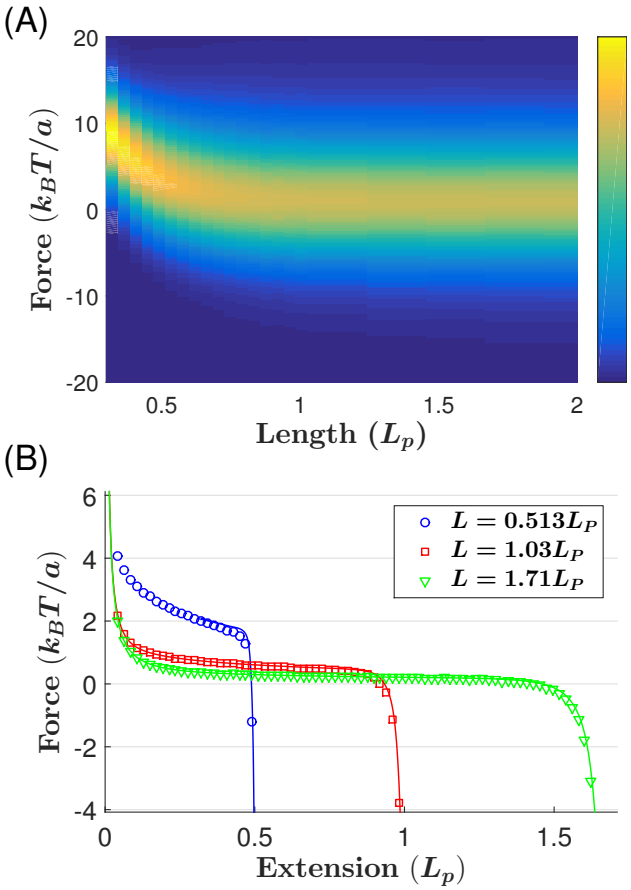

Figure 5.5: Constraint forces. (A) Distribution of forces as a function of length. The PDF values are plotted as a heat map using the colormap shown on the right. Values are at a fixed extension of 0.068 L_p. As longer chains are considered, the average force shifts towards zero, and the distribution grows broader. (B) The generalized force ($\overline{f}$) obtained from partition function (Equation 5.16), alongside the mean constraint force ($\langle\lambda\rangle$) obtained from our phase space sampling method. The two methods show good agreement over a range of extensions, though the computation of the end-to-end distribution becomes unstable for short chains at short extension.

rium statistical mechanics. Note that, as the constraints take the form of fixed distances between masses, the Fixman correction can be omitted in this instance[118].

5.6.1 Force Versus Extension

At short extensions, we see that the average force $\langle\lambda\rangle$ in the semiflexible loop is positive, indicating that the ends of the loop must be pulled inward to maintain their separation at a constant value. This result agrees with our macroscopic intuition about the force required

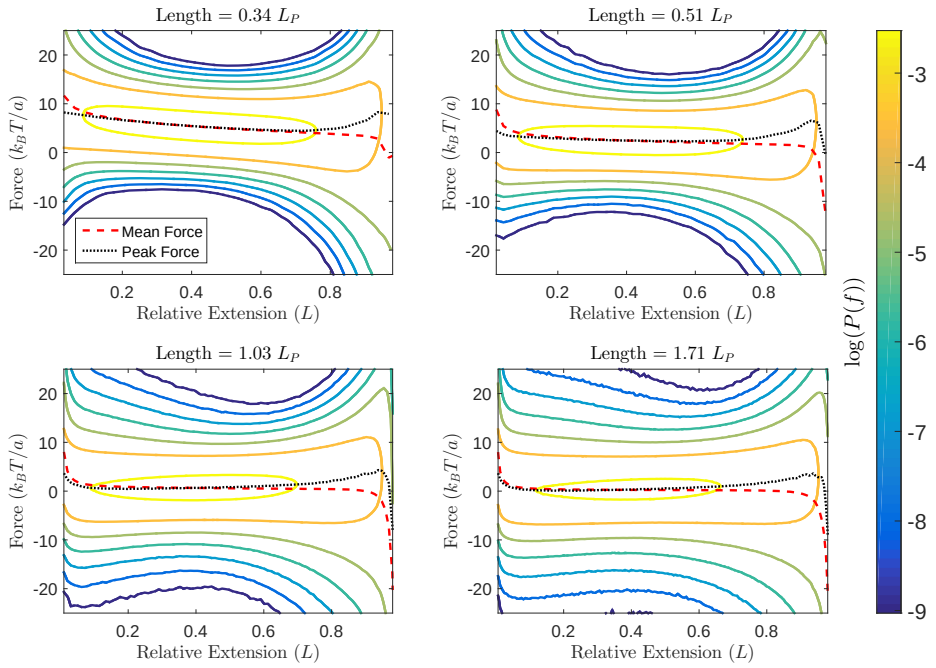

Figure 5.6: Contour plots of force distribution vs extension for four different chain lengths. Each contour line is colored according to its corresponding PDF amplitude. All figures represent data sampled for chains with a persistence length $L_p = 23.4a$. The peak value changes rapidly in regions of low and high extension, but slowly in intermediate regions. These intermediate regions also correspond to the narrowest distribution of forces.

to maintain an elastic rod in a deformed state.

As we increase extension, the average force decreases. Eventually, at large extensions, we see a strong negative force. We expect this from thermodynamic arguments, as a straight conformation is a low entropy state, and the chain will natrually want to bend away from this. Mechanically, this contractile force comes from inertial effects similar to the centrifugal force in a pendulum. In this limit, full extending lines up all such forces along the chain and produces a large effect.

5.6.2 Kinetic and Potential Contributions

We can explore the intuitions described above by looking at the decomposition into force contributions from the potential and kinetic energy. The mean force will contain contributions from an elastic force arising from internal energy stored in a deformed chain, as well

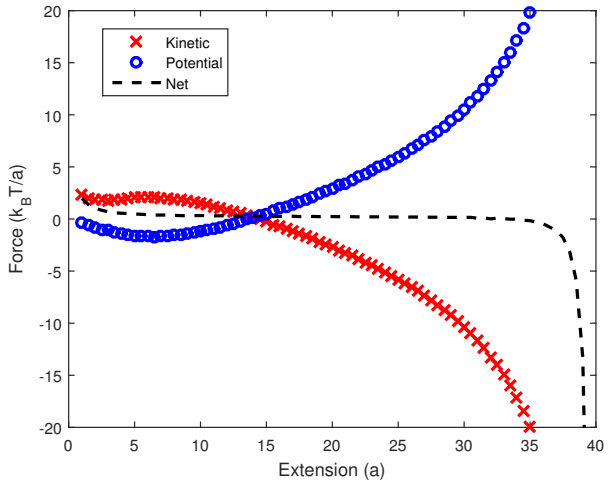

Figure 5.7: Decomposiiton of force into potential and kinetic terms as a function of chain extension. Data represent chain of legth $40a$ and persistence length $23.4a$.

as an entropic force arising from the kinetic motion of the chain[126]. If we consider only the minimum energy conformation of an elastic rod with a short fixed end-to-end distance, we see a teardrop which needs to be held with a tensile force[70].

Upon analyzing the shape-dependent potential contributions and the velocity-dependent kinetic contributions in Figure 5.7, we see it is in fact kinetic effects driving us outwards at short extension, and that the potential contributions oppose this. These two components switch signs as we sweep over the range of extensions, with the kinetic force creating a large negative result at full extensions, as our intuition regarding the entropic effects would predict.

We repeat this analysis at different contour lengths and persistence lengths, and see the mean elastic force is mostly negative, thus compressive rather than tensile (blue hollow symbols, Figure 5.8D). This negative force is compensated by a slightly larger positive entropic force (red filled symbols, Figure 5.8D) to yield a net positive mean force. Shown in Figure 5.8F are example conformations that produce positive (left) and negative (right) elastic force. The conformations with negative elastic forces typically exhibit inflection points in the contour near the ends such that the end segments bent outward exert a compressive elastic force along the end-to-end vector. Compressive elastic forces between the

ends of an elastic chain are not intuitive, but can be demonstrated even at the macroscopic level [146].

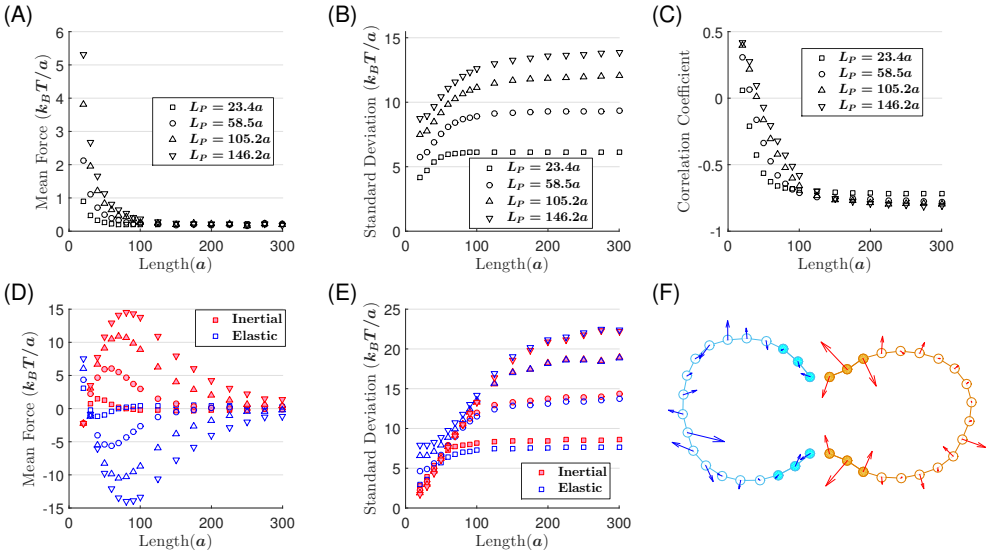

Figure 5.8: Total constraint force vs. its elastic and inertial components. (A) Average force decreases with contour length, and increases with persistence length. (B) Standard deviation increases with persistence length and contour length, plateauing around L_p. (C) Correlation coefficient decreases with length, with inertial and elastic components moving from correlated to anticorrelated. (D) Decomposition of force into potential and kinetic components versus contour length, at a constant extension of $10a$. As contour length increases, the contribution from the bending potential asymptotes to zero. The non-zero force at large contour lengths is a result of entropic contributions only. Also of note is the fact that the magnitude of both individual components generally increases with persistence length. (E) Deviation of force components follows the same trend as their sum, but reaches a larger value. (F) Two example conformations, taken from a two-dimensional ensemble with length $20a$ and end-to-end separation $4a$. Arrows indicate the instantaneous acceleration arising from the bending potential. In the conformation on the left (light blue) the first and last point are moving apart, and a linker between the two ends will be subject to a stretching force. In the conformation on the right (orange) these points are moving closer together and the linker between the two ends will be compressed. While these forces are a product of all the coordinates, the separation into stretching or compressing forces is strongly correlated with the convexity or concavity in the first three mass points from the end.

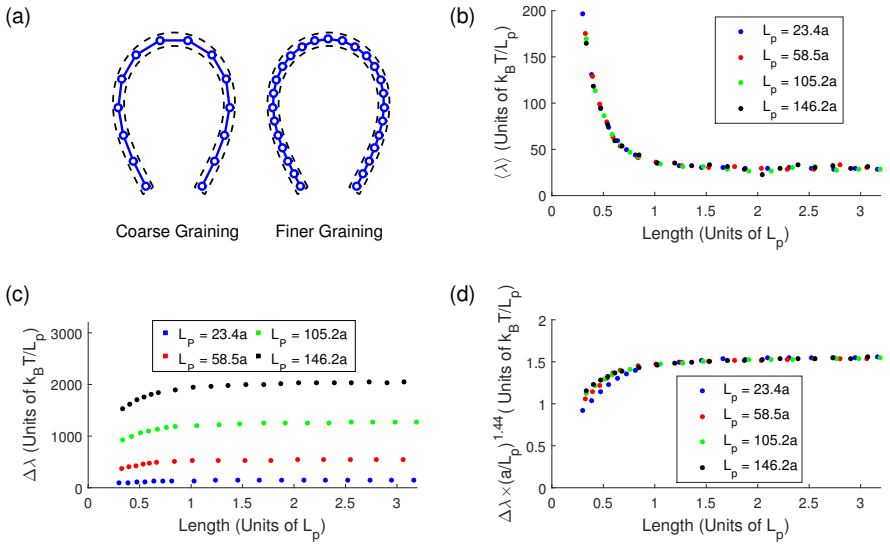

Figure 5.9: The effect of coarse-graining on force distribution. (A) Diagram of coarse-graining process. A physical polymer with a given extension and persistence length can be represented with an arbitrary number of points in a ball-and-stick model. (B) Average Force at a fixed extension $r = 0.068L_p$. As the chain increases in length, the average force decreases. The same trend, and same values, are predicted independent of number of points per persistence length. (C) Standard deviation of force at a fixed extension $r = 0.068L_p$. The fluctuation increases as the chains grow longer. Additionally, the size of fluctuations increases as more points are used to represent the chain, as each degree of freedom corresponds to more thermal energy in the system. (D) Fluctuations scaled by $(L_p/a)^{1.44}$. This measured scaling factor accounts for the growth of force fluctuations with coarse graining.

5.7 Renormalization and Scaling

As in the case of the end-pinned chain, the representation of the wormlike chain as a set of discrete links allows for some freedom in the choice in how many segments (N) are used to represent a particular length. By varying the monomer length m and length a inversely with N, and varying the bending energy directly with N, we can describe a chain with same stiffness, length, and mass density with different levels of granularity, as in Figure 5.9A. We then ask how this decision effects our reported results.

The average force $\langle\lambda\rangle$ is unnaffected by this choice in Figure 5.9B, which is unsurprising. As shown in Figure 5.5, the average force corresponds to the derivative of the free

energy, proportional to the log of the probability denisty for the wormlike chain model. As these different scalings are different representations of the same chain, they are all expected to have the same end-to-end probability density if left unconstrained. The widt of the underlying force distribution, represented by the standard deviation $\Delta\lambda$, increases with the number of monomers N regardless of the chain length. An approximate scaling law exists between this fluctuation and the chain granularity, where $\Delta\lambda \sim N^{1.44}$. Rescaling by this factor collapses the fluctuation onto a single curve in Figure 5.9D. While the amplitude of fluctuations maintains some dependence on the coarse-graining, the shape of the curve appears to be universal. The growing fluctuation may appear to imply that force fluctuation, in contrast to average force, increases with the degrees of freedom with no bound, similar to a Casimir-like force between two plates[100]. However, the details of the polymer chain at the monomeric level will place one absolute bound on how fine we can realistically proceed. Additionally, while we assume the equipartition theorem can be applied to all the modal coordinates, the independence of modes close in space may be limited by the viscosity of the surrounding solvent.

5.8 Toy Model

The observed scaling of force fluctuations in our system can be explained with the introduction of a toy model that shares many of the features of interest. We consider a beam of mass M stretched at length L, immersed in a heat bath. We coarse-grain it into N point masses m connected by $N + 1$ springs of stiffness κ (Figure 5.10), each with zero equilibrium extension. This toy model represents a simplification of that employed by others [147, 148], wherein each spring has a non-zero equilibrium extension. The beginning of the first spring is fixed at the origin, and the end of the last spring is held at the point $L\hat{e}_x$ in Cartesian space. In this model, all bonds are extensible, and thus the force of constraint is entirely localized to the last spring in the system, with no dependence on the velocity of

the point masses. The Hamiltonian is easily separable into kinetic and potential terms,

$$\mathcal{H} = \sum_{i=1}^{N} \frac{\|\mathbf{p}_i^2\|}{2m} + \frac{\kappa\|\mathbf{x}_1\|^2}{2} + \sum_{i=2}^{N} \frac{\kappa\|\mathbf{x}_i - \mathbf{x}_{i-1}\|^2}{2} + \frac{\kappa\|\mathbf{x}_N - \mathbf{L}\|^2}{2} \tag{5.17}$$

where each $\mathbf{p}_i$ and $\mathbf{x}_i$ represents three cartesian components of momentum or position vectors, and $\mathbf{L} = L\mathbf{e}_x$ is the displacement vector between the first and last oscillator. This separability allows us to use analytical means to derive force fluctuations.

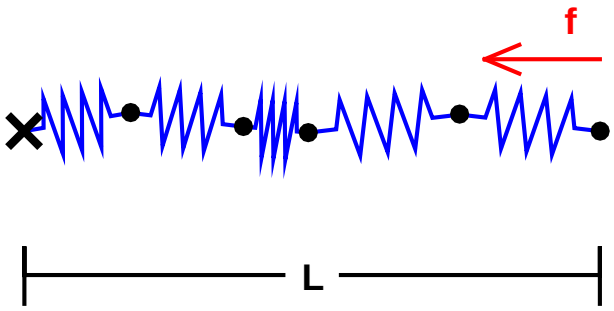

Figure 5.10: Schematic of toy model. Each of the springs has a relaxed length of zero, and they are held at a constant extension of L. Representing the same system with more springs requires that each one be stiffer to maintain the same average force. The masses are free to move in three dimensions.

5.8.1 Force Fluctuations

The total partition function for the system can be written down as

$$Z = \int\int \exp(-\beta\mathcal{H}) d^{3N}p \, d^{3N}x \tag{5.18}$$

All momentum integrands are Gaussians, simplifying our partition function to

$$Z = \left(\frac{2\pi m}{\beta}\right)^{3N/2} \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} e^{-\frac{\beta\kappa}{2}(\|\mathbf{x}_1\|^2 + \|\mathbf{x}_2 - \mathbf{x}_1\|^2 \cdots + \|\mathbf{x}_N - \mathbf{L}\|^2)} d\mathbf{x}_1 \ldots d\mathbf{x}_N \tag{5.19}$$

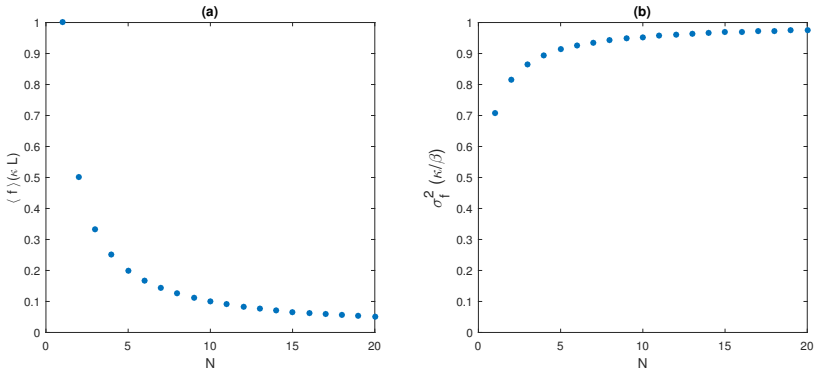

Figure 5.11: (a) Mean force in toy model. More springs relieve the stress in the system, similar to the effect of lengthening the contour in the isometric ensemble. (b) Fluctuation in toy model. More degrees of freedom means large fluctuations, even as average force goes down.

Each position integral can be carried out in turn on each Cartesian position, producing

$$Z = \frac{1}{(N+1)^{3/2}} \left(\frac{2\pi}{\beta} \sqrt{\frac{m}{\kappa}} \right)^{3N} e^{-\beta\kappa L^2/2(N+1)} \tag{5.20}$$

which is identical to the partition function of a Gaussian chain[148]. The ensemble average of the instantaneous force ($\kappa(x_N - L)$) along $\mathbf{e}_x$, is then related to the derivative of the partition function,

$$\bar{f} = \frac{\int_{-\infty}^{\infty} \kappa(x_N - L) e^{-\beta U(\mathbf{x})} d^{3N}x}{\int_{-\infty}^{\infty} e^{-\beta U(\mathbf{x})} d^{3N}x} = \frac{\int_{-\infty}^{\infty} -\frac{1}{\beta}\frac{\partial}{\partial L} e^{-\beta U(\mathbf{x})} d^{3N}x}{\int_{-\infty}^{\infty} e^{-\beta U(\mathbf{x})} d^{3N}x} = -\frac{1}{\beta}\frac{d \ln Z}{dL}. \tag{5.21}$$

Hence, using Equation 5.20, we obtain the generalized force conjugate to L

$$\bar{f} = \frac{\kappa}{N+1} L. \tag{5.22}$$

This equation is simply Hooke's law with bulk stiffness $k = \frac{\kappa}{N+1}$, analogous to the persistence length for the wormlike chain case. Similar to Figure 5.9b, we see in Figure 5.11(a) that increasing N relieves the average stress on the system analogous to increasing the contour length. The fluctuation can similarly be obtained by taking the second order derivative

of Z,

$$\sigma_f^2 = \overline{f^2} - \overline{f}^2 = -\frac{1}{\beta^2}\frac{d^2 \ln Z}{dL^2} + \frac{\kappa}{\beta} \tag{5.23}$$

which produces

$$\sigma_f = \sqrt{\frac{\kappa N}{\beta(N+1)}} \approx \sqrt{\frac{kN}{\beta}} \propto N^{0.5}. \tag{5.24}$$

When considered at constant κ, this fluctuation agrees qualitatively with the growth and saturation as a function of N displayed in Figure 5.9c. (Figure 5.11 (b)) Additionally, at fixed bulk extensibility k, we again see the phenomena where the amplitude of force fluctuations increases with degrees of freedom added to the system even when the average force is assumed not to, as in Figure 5.9d. However, the scaling law is different, likely as a result of the way these degrees of freedom are coupled to the force. In our toy model, the force is only a function of the extension of the last spring, whereas in the discretized wormlike chain, it is a complicated function that may depend on all the generalized position and momentum coordinates.

5.8.2 Work Fluctuations

What is the impact of this fluctuating force? In statistical mechanics of many particle systems, fluctuation of intensive parameters such as force still appears to be a subject of discussion[104, 149, 150]. In our example of a single polymer chain, the fluctuating force can be given a mechanistic interpretation in terms of the actual work transmissible during a short period of time. Here, using the same Gaussian chain model above, we show that the change in energy during an adiabatic extension of the chain is bounded with respect to N, despite the unbounded fluctuations in the force. Imagine that the chain is allowed to extend by Δ over a time period τ, shorter than the characteristic collision time between the chain and molecules in the surrounding heat bath. Using the initial microstate of the chain, we can calculate the energy difference as a result of this extension. Equations of motion in the y and z dimensions will be separable from those in the x dimension, and will not

contribute to the force. The equations of motion in the x dimension for N oscillators can be represented as a matrix equation

$$\frac{m}{\kappa}\begin{pmatrix} \ddot{x}_1 \\ \ddot{x}_2 \\ \vdots \\ \ddot{x}_{N-1} \\ \ddot{x}_N \end{pmatrix} = -\begin{pmatrix} 2 & -1 & 0 & \cdots & 0 \\ -1 & 2 & \ddots & \ddots & \vdots \\ 0 & \ddots & \ddots & \ddots & 0 \\ \vdots & \ddots & \ddots & 2 & -1 \\ 0 & \cdots & 0 & -1 & 2 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_{N-1} \\ x_N \end{pmatrix} + \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ -L(t) \end{pmatrix} \tag{5.25}$$

with a tri-diagonal matrix relating x_i and $\ddot{x}_i$, accompanied by an inhomegenous vector representing the overall extension of the system. Solutions to this system will thus take the form

$$x_i = \sum_j^N A_{ij}\cos(\omega_j t) + B_{ij}\sin(\omega_j t) + C_i + D_i t \tag{5.26}$$

where the summation part of the expression satisfies the homogeneous part of the equation, and the linear terms satisfy the inhomogeneous component. In the homogeneous solution, the frequencies ω_j^2 will correspond to the eigenvalues of the matrix.

$$\omega_j = 2\sqrt{\frac{\kappa}{m}}\sin\left(\frac{1}{2}\frac{\pi j}{N+1}\right) \tag{5.27}$$

Note that while κ scales as $(N+1)k$, the mass of each point will scale as $m = M/(N+1)$ to maintain a fixed linear mass density. ω_j then scales as $N+1$ when we consider finer graining of the system. These frequencies will be unchanged by the extension of the system, which is confined to the inhomogenous part of the equation. Representing the extension by a function $L(t)$ that is piecewise linear in time, increasing uniformly from L to $L+\Delta$ in the interval $t = 0$ to $t = \tau$, we can obtain matching conditions for the coefficients $\mathbf{A}, \mathbf{B}, \mathbf{C}$ and $\mathbf{D}$ at times $t = 0$ and $t = \tau$, when the extension begins and ends. This is done by assuming continuity of all position and momentum coordinates. Computing the difference between the Hamiltonian before and after the extension allows us to find the work done.

Assuming an equilibrium distribution of energy among the normal modes, we can find the expected work and its fluctuation (see Appendix for detailed calculation.). Evaluating the expectation of this difference reveals the average work

$$\overline{W} = 4k\Delta^2 \sum_{i=1}^{N} \frac{\sin^2(\omega_i \tau/2)}{\omega_i^2 \tau^2} \cos^2\left(\frac{1}{2}\frac{i\pi}{N+1}\right) + \frac{k}{2}(2L\Delta + \Delta^2) \tag{5.28}$$

Taking the lowest order terms in Δ, we find $\overline{W} \approx kL\Delta = \Delta\overline{f}$, using the expression for $\overline{f}$ in Equation 5.22. The average value of work squared can also be found, and used to obtain the fluctuation.

$$\overline{W^2} - \overline{W}^2 = \frac{8k\Delta^2}{\beta} \sum_{i=1}^{N} \frac{\sin^2(\omega_i \tau/2)}{\omega_i^2 \tau^2} \cos^2\left(\frac{1}{2}\frac{i\pi}{N+1}\right) \tag{5.29}$$

If we take the limit of small τ, independent of N, we can use the approximation $\sin(\omega_i\tau/2) \approx \omega_i\tau/2$ to reduce this to

$$\sigma_W^2 = \overline{W^2} - \overline{W}^2 \approx \frac{2k\Delta^2}{\beta} \sum_{i=1}^{N} \cos^2\left(\frac{i\pi}{2(N+1)}\right) = \frac{kN\Delta^2}{\beta} \tag{5.30}$$

which is the same result we found using equilibrium statistical mechanics (Equation 5.24), albeit in terms of $\sigma_W = \sigma_F \times \Delta$ However, for any finite τ, we will reach a limit of rescaling where $\omega_i\tau$ is no longer negligible. This will provide a cap on our growing fluctuations, and define a scale of coarse-graining below which further fineness will not produce any change in the results. With some effort, the expression in Equation 5.29 can be manipulated to reveal

$$\sigma_W^2 = \frac{8k\Delta^2}{\beta}((N+1)(2J_0(x) + \pi J_1(x)H_0(x) - \pi J_0(x)H_1(x) - \frac{1}{2} + \frac{M(J_0(x)-1))}{2(N+1)\tau^2 k} - \frac{J_1(x)}{\tau\sqrt{k/M}}) \tag{5.31}$$

where J_i are the Bessel functions of the first kind, and H_i are the Struve functions. All of these take the same argument $x = 2(N+1)\tau\sqrt{k/M}$. Taking the limit or large N, the two terms on the right will become negligible, leaving us with

$$\sigma_W^2 = \frac{8k\Delta^2}{\beta}\left(\frac{\sqrt{M}}{\tau\sqrt{k}} - \frac{1}{2}\right) \tag{5.32}$$

for $M/\tau^2 k > 1/4$. The fluctuations are limited in N, but can increase if the time of the extension is short enough.

5.9 Summary

By introducing a new set of coordinates, we are able to sample the isometric ensemble in both position and momentum space. This distribution in phase space, combined with Lagrangian mechanics, gives us a distribution of forces representing semiflexible polymers held at fixed extension. We can see that, on average, this result aligns with the generalized force derived from the free energy. This agreement occurs in spite of some distinctions between the ensemble used to obtain the partition function [75], which consists purely of spatial conformations of a continuous chain, versus the discrete chain constrained in both position and momentum coordinates used for our analysis. The differences between conditional and constrained averages have studied extensively in the literature[151, 152, 118, 153], but are not expected to have an effect in this instance[118].

Entropic forces are typically understood by counting the number of static conformations[154, 155]. Here, we demonstrate a mechanistic origin of this phenomena, arising from thermally-excited motions of the polymer chain. This kinetic origin of entropic forces had been described before for other, simpler systems[99, 156, 157, 158].Upon decomposing this force into contributions from the elasticity of the chain and from its thermal motion, we find unexpectedly that the inertial contribution from the movement of the chain is driving it toward larger extenions, but that the elastic force is in fact pushing the ends inward.

This runs counter to our intuition about elastic deformation at the macroscopic scale.

After establishing the success of phase-space sampling for a geometry where the partition function is obtainable, we can extend the method to situations where it is not so easily computed. Examples include DNA loops with intrinsic sequence-dependent shape and flexibility. This could also be extended to compute torques conjugate to constrained angles, similar to forces conjugate to constrained distances. Information on the full distribution of forces may prove useful for predicting lifetimes in configurations that can be destabilized by critical forces crossing some threshold value.

CHAPTER 6
CONCLUSIONS

While initially this work was motivated by questions about loop formation and breaking in DNA strands, it has led us to develop methods applicable to a broad range of systems in polymer physics. We focused on finding a way to explore the ensemble of states in both position and momentum space, so as to produce the full distribution of forces for a given ensemble. This approach can reveal new information about the kinetics of these systems, where processes such as loop-breaking or the severing of surface-attached chains may depend on fluctuations in the constraining forces.

6.1 Mechanical Force and Thermodynamic Force

Sampling the distribution of states in phase-space can be done through multiple means. We find an efficient path by applying a pseudo-potential corresponding to the size of the momentum space[153] and sampling momentum values once position coordinates have been obtained[35]. After obtaining both sets of coordinates, we can use classical mechanics to find a constraint force along coordinates of interest. This gives a clear mechanistic origin for the entropic force, arising from inertial forces in the thermally-excited links of the polymer chain. Even in the absence of a potential energy, we see this emerge as a fluctuation-induced quantity similar similar to Casimir force [100] and depletion force[159]. The average of these mechanical forces lines up with the mean force predicted from the partition function in the absence of a bounding surface.

6.1.1 Effect of Confinement

Confinement of polymer chains restricts the accessible conformational space, and so the chains will have a tendency to break free as they move towards higher entropy states. We

can measure this change in entropy with respect to a given coordinate by simply counting the fraction of conformations that are compatible with the boundary at different values of the coordinate. This is a slow process for multiple coordinates, but is a simple way of finding a mean force. We see this thermodynamic mean force arises when this confinement is applied in such a way as to bias the set of shapes the chain can take on, skewing our sampled mechanical forces towards larger values and altering the average.

6.1.2 Difference Between Forces

In many instances, the mean of the sampled constraint force matches the result predicted by thermodynamic arguments. The term relating the constrained and unconstrained ensembles, the so-called Fixman correction[117], does not appear in cases where all the constrained are of the form of fixed distances between mass points. When confinement is introduced, not by fixing one coordinate but by limiting the range a particular coordinate can explore, this correction will be non-zero, but generally small. Instead, the chief source of discrepancy can also be identified as having a mechanistic origin. In this case, it is the collisions of the polymer against the confining wall. These represent a set of lower dimensionality, and as such are not covered by phase space sampling, but each collision produces a finite impulse, and as a result their contribution to the average force is non-negligible.

6.2 Scaling of Forces

For the case of free or semi-flexible polymer chains fixed in space at one end, we find a uniform force of $2k_BT/a$ along the chain. This occurs independent of any other parameters, such as the length, flexibility, or bond index. This result can be predicted from simple arguments. Such scaling links it inextricably to the level of discretization chosen for the system- the more relevant quantity to consider for physical problems may be the increase in this force when the chain is subject to confinement or constraint, as well as the breadth of the fluctuation in this force.

In the case of the isometric ensemble, we find that the average force scales inversely with the persistence length. This is what we would expect, as it lines up with the thermodynamic mean force predicted from a continuous wormlike chain. As this prediction involves no inherent scale of the discretization of the chain, it is as we would expect that the actual force remains fixed- the numerical value decreases as the units increase.

For end-pinned chains confined by a surface, we find a scaling law for the force measured from the partition function. Contrary to our expectation, this scaling law goes as $L_P^{-0.7}$, rather than inversely with the persistence length. Such behavior would allow us to identify the coarse-grained result with a single value, independent of the scale, as the units of force scale proportionally to the persistence length. This result may arise from the discontinuity in the potential at an infinitely hard surface, or it may be the result of some other feature lost in our coarse-graining process.

In contrast, no clear scaling law emerges in the case of sampled constraint forces in this instance. The two measurements of force disagree as a result of collisions with the surface. This would in turn suggest that the average recoil along the chain scales with persistence length in the same manner as the mean force measured from the partition function, but with the opposite sign. Further analysis may be required to verify that this is indeed the case.

6.3 Scaling of Fluctuations

We find in the case of end-pinned chains that the width of the force distribution, characterized by the standard deviation, generates the same shape when plotted as a function of bond index along the length of the curve. The amplitude of this curve scales with the persistence length as $L_P^{1.74}$. The growth is to be expected, as a finer-grained representation of the system possess more degrees of freedom to fluctuate in. Physically, we expect this growth to be capped at some scale, by the minimum scale of discretization in the polymer or by the minimal scale excitable by the solution the polymer resides in.

For the isometric ensemble, the fluctuations scale as $L_P^{1.44}$. This positive exponent

makes it subject ot many of the same considerations as in the case of end-pinned chains, where a limiting scale is necessary to cap increasing fluctuation. The smaller power in the scaling law suggests that added degrees of freedom do not couple to the end-to-end constraint as strongly as to the bond length constraints.

6.4 Future Directions

6.4.1 Boundary Conditions

As an initial study, our model of the isometric ensemble makes very few assumptions about the environment apart from the polymer chain. The absence of any constraints on the ends of the chain may be an accurate representation for a short, looped DNA strand free in solution. It may also be interesting to investigate the effect of a bending potential applied at the ends of the strand. Such a case might arise in a DNA sequence bound by a protein complex, where the shape of the protein imposes additional constraints on the angles at the ends of the loop[160]. We may find that this alters the ensemble of states in a way that helps to stabilize the loop.

6.4.2 Detailed Polymer Models

Specific applications of the phase-space sampling approach to DNA could also be found, by adding details particular to the DNA molecule. We could even increase utility of the methods deveoped here by applying them to instances where the distribution of end-to-end states is unknown. To obtain an average force from the partition function in chapter five, we made use a theoretical model assuming uniform flexibility and no intrinsic curvature along the length of the polymer chain. Finding the analogous distribution after introducing sequence-specific curvature and flexibility may prove challenging- by applying a phase-space sampling approach, we can compute the derivative of the unknown free energy by way of the average force, at the same time as finding the distribution of forces.

6.4.3 Physical Solutions

The majority of this work has been predicated on the assumption that all the degrees of freedom can be perturbed independently. In an overdamped system, we expect the motion of particles to be correlated with the local fluid velocity. A coarse-grained or simplified model of a solution, such as a lattice-Boltzmann fluid[161], could be used to represent the environment around a looped or surface-bound polymer chain. This could then offer an alternative way to sample the space of velocities, that incorporated realistic parameters such as viscosity by way of the fluid model.

Appendices

APPENDIX A

SCALING LAWS FOR A MIDDLE-PINNED CHAIN

To obtain a prediction for the RMS end-to-end distance (r_{rms}) for the middle pinned case in the flexible limit, we consider two one-dimensional random walks with unit step size on the lattice starting at the origin with a perfectly absorbing boundary. Both random walks will be confined to the same half-space. For a single random walk of total N steps, we define the number of right steps as n_R, and the number of left steps as n_L. The end-to-end distance is thus $r = n_R - n_L$. We seek a general solution that is valid for any N using binomial coefficients. For $N \to \infty$, one can use the continuous probability distribution to obtain the solution more easily. The total number of conformations with this r in the absence of the boundary is:

$$C(N; n_R) = \binom{N}{n_R} \tag{A.1}$$

From the reflection principle, we can obtain the number of walks in the presence of the boundary

$$C'(N; n_R) = \binom{N}{n_R} - \binom{N}{n_R + 1} \tag{A.2}$$

Values of n_R less than $N/2 (\equiv M_N)$ will not be allowed. Summing over all lengths gives us the total number of walks in the right half space.

$$C'(N) = \sum_{n_R = M_N}^{N} C'(N; n_R) = \binom{N}{N/2} \tag{A.3}$$

We can now consider our middle pinned chain as two such random walks of N steps each, and compute the mean square distance between their respective ends.

$$r_{rms}^2 = \frac{1}{C'(N)^2} \sum_{n_{R1} = M_N}^{N} C'(N; n_{R1}) \sum_{n_{R2} = M_N}^{N} C'(N; n_{R2}) 4(n_{R1} - n_{R2})^2 \tag{A.4}$$

Expanding the product at the right gives

$$r_{rms}^2 = \frac{4}{C'(N)^2} \sum_{n_{R1}=M_N}^{N} C'(N;n_{R1}) \sum_{n_{R2}=M_N}^{N} C'(N;n_{R2})(n_{R1}^2 - 2n_{R1}n_{R2} + n_{R2}^2)$$

Recognizing that the n_{R1}^2 and n_{R2}^2 terms are interchangeable, this further simplifies to

$$r_{rms}^2 = 8\left(\frac{1}{C'(N)} \sum_{n_R=M_N}^{N} C'(N;n_R)n_R^2 - \left(\frac{1}{C'(N)} \sum_{n_R=M_N}^{N} C'(N;n_R)n_R\right)^2\right)$$

which we recognize as simply

$$r_{rms}^2 = 8\left(\langle n_R^2\rangle - \langle n_R\rangle^2\right) \tag{A.5}$$

The second term can be found by expanding

$$\sum_{n_R=M_N}^{N} C'(N,n_R)n_R = \sum_{n_R=M_N}^{N} C'(N-1,n_R-1)N + \binom{N}{n_R+1}$$

When we add over the first two terms in the summand, all but one will cancel, leaving us with

$$\sum_{n_R=M_N}^{N} C'(N,n_R)n_R = \binom{N}{M_N}M_N + \sum_{n_R=M_N}^{N} \binom{N}{n_R+1}$$

The remaining sum can be eliminated by making use of the fact that the binomial coefficients $\binom{N}{n_R}$ sum to 2^N when n_R runs from 0 to N. Dividing by $C'(N)$ will leave us with

$$\langle n_R\rangle = \left(\frac{N-1}{2}\right) + \frac{2^{N-1}}{\binom{N}{M_N}} \tag{A.6}$$

The first term can be found by rewriting the sum as

$$\sum_{n_R=M_N}^{N} C'(N;n_R)n_R^2 = \sum_{n_R=M_N}^{N} C'(N-2;n_R-2)(N^2-N) + \left[\binom{N-1}{n_R-1} + \binom{N-1}{n_R}\right] N - \binom{N}{n_R+1} \quad \text{(A.7)}$$

which can be simplified in a similar manner to that used for $\langle n_R \rangle$. This results in a term

$$\langle n_R^2 \rangle = \left(\frac{N^2+2}{4}\right) + \frac{(N-1)2^{N-1}}{\binom{N}{M_N}} \qquad N = \text{even} \qquad \text{(A.8a)}$$

$$\langle n_R^2 \rangle = \left(\frac{N^2+3}{4}\right) + \frac{(N-1)2^{N-1}}{\binom{N}{M_N}} \qquad N = \text{odd} \qquad \text{(A.8b)}$$

Substitution into the expression for r_{rms}^2 produces, after some cancellation,

$$r_{rms}^2 = 4N + 2 - \frac{2^{2N+1}}{\binom{N}{M_N}^2} \qquad N = \text{even} \qquad \text{(A.9a)}$$

$$r_{rms}^2 = 4N + 4 - \frac{2^{2N+1}}{\binom{N}{M_N}^2} \qquad N = \text{odd} \qquad \text{(A.9b)}$$

Factoring out the total number of steps $2N$ gives us, in the limit of large N

$$r_{rms}^2 = \left(2 - \frac{2^{2N}}{N\binom{N}{N/2}^2}\right) 2N$$

Now we make use of the fact that the binomial distribution converges to a Gaussian distribution in the limit of large N. This gives us a value

$$\lim_{N\to\infty} \frac{1}{2^N}\binom{N}{N/2} = \sqrt{\frac{2}{\pi N}}$$

Substituting back in yields

$$r_{rms}^2 = \left(2 - \frac{\pi}{2}\right) 2N \qquad \text{(A.10)}$$

As in the case of end-pinned chains, the mean square displacement in other, unconstrained, dimensions will be unaffected. In the limit of large N, the overall mean squared distance will be the mean of the mean squared distances for each dimension. In three dimensions, this is

$$r_{rms}^2 = \frac{1}{3}\left(4 - \frac{\pi}{2}\right) N \tag{A.11}$$

Note that this only applies for large N; for shorter random walks, the steps are more likely to be along the unconstrained dimensions than the constrained dimension and the RMS distance will be closer to N.

REFERENCES

[1] R. Schleif, "Dna looping," *Annual review of biochemistry*, vol. 61, no. 1, pp. 199–223, 1992.

[2] J. S. Williams, T. T. Eckdahl, and J. N. Anderson, "Bent dna functions as a replication enhancer in saccharomyces cerevisiae.," *Molecular and cellular biology*, vol. 8, no. 7, pp. 2763–2769, 1988.

[3] R. R. Koepsel and S. A. Khan, "Static and initiator protein-enhanced bending of dna at a replication origin," *Science*, vol. 233, no. 4770, pp. 1316–1318, 1986.

[4] H. Müller and H. E. Varmus, "Dna bending creates favored sites for retroviral integration: An explanation for preferred insertion sites in nucleosomes.," *The EMBO journal*, vol. 13, no. 19, p. 4704, 1994.

[5] A. J. Andrews and K. Luger, "Nucleosome structure (s) and stability: Variations on a theme," *Annual review of biophysics*, vol. 40, pp. 99–117, 2011.

[6] C. Mao, W. Sun, N. C. Seeman, *et al.*, "Assembly of borromean rings from dna," *Nature*, vol. 386, no. 6621, pp. 137–138, 1997.

[7] J. Chen and N. C. Seeman, "Synthesis from dna of a molecule with the connectivity of a cube.," *Nature*, vol. 350, no. 6319, pp. 631–633, 1991.

[8] J. Bath, S. J. Green, and A. J. Turberfield, "A free-running dna motor powered by a nicking enzyme," *Angewandte Chemie*, vol. 117, no. 28, pp. 4432–4435, 2005.

[9] S. F. Wickham, J. Bath, Y. Katsuda, M. Endo, K. Hidaka, H. Sugiyama, and A. J. Turberfield, "A dna-based molecular motor that can navigate a network of tracks," *Nature nanotechnology*, vol. 7, no. 3, pp. 169–173, 2012.

[10] J. Liphardt, B. Onoa, S. B. Smith, I. Tinoco, and C. Bustamante, "Reversible unfolding of single rna molecules by mechanical force," *Science*, vol. 292, no. 5517, pp. 733–737, 2001.

[11] T. Harmer, M. Wu, and R. Schleif, "The role of rigidity in dna looping-unlooping by arac," *Proceedings of the National Academy of Sciences*, vol. 98, no. 2, pp. 427–431, 2001.

[12] S Cuesta-Lopez, J Errami, F Falo, and M Peyrard, "Can we model dna at the mesoscale?" *Journal of biological physics*, vol. 31, no. 3-4, pp. 273–301, 2005.

[13] P. H. Weigel, "Functional characteristics and catalytic mechanisms of the bacterial hyaluronan synthases," *IUBMB life*, vol. 54, no. 4, pp. 201–211, 2002.

[14] K. H. Iwasa and A. M. Florescu, "A molecule that detects the length of dna by using chain fluctuation," *ArXiv preprint arXiv:1510.00011*, 2015.

[15] E. F. Koslover and A. J. Spakowitz, "Force fluctuations impact kinetics of biomolecular systems," *Physical Review E*, vol. 86, no. 1, p. 011 906, 2012.

[16] L. Boltzmann and S. G. Brush, *Vorlesungen über Gastheorie*. Akademische Druck-u. Verlagsanstalt, 1981, vol. 1.

[17] O Kratky and G Porod, "Röntgenuntersuchung gelöster fadenmoleküle," *Recueil des Travaux Chimiques des Pays-Bas*, vol. 68, no. 12, pp. 1106–1122, 1949.

[18] N. Saitô, K. Takahashi, and Y. Yunoki, "The statistical mechanical theory of stiff chains," *Journal of the Physical Society of Japan*, vol. 22, no. 1, pp. 219–226, 1967.

[19] J. G. Elias and D. Eden, "Transient electric birefringence study of the persistence length and electrical polarizability of restriction fragments of dna," *Macromolecules*, vol. 14, no. 2, pp. 410–419, 1981.

[20] J. Bednar, P. Furrer, V. Katritch, A. Stasiak, J. Dubochet, and A. Stasiak, "Determination of dna persistence length by cryo-electron microscopy. separation of the static and dynamic contributions to the apparent persistence length of dna," *Journal of molecular biology*, vol. 254, no. 4, pp. 579–594, 1995.

[21] C Bouchiat, M. Wang, J.-F. Allemand, T Strick, S. Block, and V Croquette, "Estimating the persistence length of a worm-like chain molecule from force-extension measurements," *Biophysical journal*, vol. 76, no. 1, pp. 409–413, 1999.

[22] S. B. Smith, L. Finzi, and C. Bustamante, "Direct mechanical measurements of the elasticity of single dna molecules by using magnetic beads," *Science*, vol. 258, no. 5085, pp. 1122–1126, 1992.

[23] M. D. Wang, H. Yin, R. Landick, J. Gelles, and S. M. Block, "Stretching dna with optical tweezers.," *Biophysical journal*, vol. 72, no. 3, p. 1335, 1997.

[24] J. Shimada and H. Yamakawa, "Ring-closure probabilities for twisted wormlike chains. application to DNA," *Macromolecules*, vol. 17, no. 4, pp. 689–698, 1984.

[25] D Shore, J Langowski, and R. L. Baldwin, "DNA flexibility studied by covalent closure of short fragments into circles.," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 78, no. 8, pp. 4833–4837, Aug. 1981, PMID: 6272277 PMCID: PMC320266.

[26] S. B. Smith, Y. Cui, and C. Bustamante, "Overstretching b-dna: The elastic response of individual double-stranded and single-stranded dna molecules," *Science*, vol. 271, no. 5250, pp. 795–799, 1996.

[27] T. Norisuye and H. Fujita, "Excluded-volume effects in dilute polymer solutions. xiii. effects of chain stiffness," *Polymer Journal*, vol. 14, no. 2, pp. 143–147, 1982.

[28] K. F. Freed, "Wiener integrals and models of stiff polymer chains," *The Journal of Chemical Physics*, vol. 54, no. 4, pp. 1453–1463, 1971.

[29] H. Yamakawa, "Statistical mechanics of wormlike chains: Path integral and diagram methods," *The Journal of Chemical Physics*, vol. 59, no. 7, pp. 3811–3815, 1973.

[30] E. Guth and H. M. James, "Elastic and thermoelastic properties of rubber like materials," *Industrial & Engineering Chemistry*, vol. 33, no. 5, pp. 624–629, 1941.

[31] F Bueche, "Physical properties of polymers (interscience, new york, 1962)," vol. 1, p. 37, 1973.

[32] P. L. Taylor and J. Tabachnik, "Entropic forcesmaking the connection between mechanics and thermodynamics in an exactly soluble model," *European Journal of Physics*, vol. 34, no. 3, p. 729, 2013.

[33] S. Mehraeen and A. J. Spakowitz, "Intrinsic fluctuations lead to broad range of transduced forces in tethered-bead single-molecule experiments," *Physical Review E*, vol. 86, no. 2, p. 021 902, 2012.

[34] T.-L. Kuo, S. Garcia-Manyes, J. Li, I. Barel, H. Lu, B. J. Berne, M. Urbakh, J. Klafter, and J. M. Fernández, "Probing static disorder in arrhenius kinetics by single-molecule force spectroscopy," *Proceedings of the National Academy of Sciences*, vol. 107, no. 25, pp. 11 336–11 340, 2010.

[35] "Equipartition principle for internal coordinate molecular dynamics.," *Journal of chemical theory and computation*, vol. 8, no. 8, pp. 2581–2587,

[36] W. Den Otter and W. Briels, "Free energy from molecular dynamics with multiple constraints," *Molecular Physics*, vol. 98, no. 12, pp. 773–781, 2000.

[37] E. Darve and A. Pohorille, "Calculating free energies using average force," *The Journal of Chemical Physics*, vol. 115, no. 20, pp. 9169–9183, 2001.

[38] T. T. Le and H. D. Kim, "Measuring shape-dependent looping probability of DNA," *Biophysical Journal*, vol. 104, no. 9, pp. 2068–2076, May 2013.

[39] R. Vafabakhsh and T. Ha, "Extreme bendability of DNA less than 100 base pairs long revealed by single-molecule cyclization," *Science*, vol. 337, no. 6098, pp. 1097–1101, Aug. 2012.

[40] P. Cifra, Z. Benkova, and T. Bleha, "Persistence lengths and structure factors of wormlike polymers under confinement," *The Journal of Physical Chemistry B*, vol. 112, no. 5, pp. 1367–1375, Feb. 2008.

[41] A. Silberberg, "The adsorption of flexible macromolecules part i. the isolated macromolecule at a plane interface," *The Journal of Physical Chemistry*, vol. 66, no. 10, pp. 1872–1883, Oct. 1962.

[42] E. A. DiMarzio and F. L. McCrackin, "OneDimensional model of polymer adsorption," *The Journal of Chemical Physics*, vol. 43, no. 2, pp. 539–547, Jul. 1965.

[43] "Conformation of an isolated polymer molecule at an interface," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 53, no. 1,

[44] K. Motomura and R. Matuura, "Conformation of adsorbed polymeric chain. II," *The Journal of Chemical Physics*, vol. 50, no. 3, pp. 1281–1287, Feb. 1969.

[45] A. Takahashi and M. Kawaguchi, "The structure of macromolecules adsorbed on interfaces," in *Behavior of Macromolecules*, ser. Advances in Polymer Science 46, Springer Berlin Heidelberg, Jan. 1982, pp. 1–65, ISBN: 978-3-540-11640-0, 978-3-540-39453-2.

[46] O. J. Hehmeyer, G. Arya, A. Z. Panagiotopoulos, and I. Szleifer, "Monte carlo simulation and molecular theory of tethered polyelectrolytes," *The Journal of Chemical Physics*, vol. 126, no. 24, pp. 244 902–244902–11, Jun. 2007.

[47] B Zhao and W. Brittain, "Polymer brushes: Surface-immobilized macromolecules," *Progress in Polymer Science*, vol. 25, no. 5, pp. 677–710, Jun. 2000.

[48] E. Lindberg and C. Elvingson, "Monte carlo simulation of polymer brushes attached to a spherical surface," *The Journal of Chemical Physics*, vol. 114, no. 14, pp. 6343–6352, Apr. 2001.

[49] R. Toral and A. Chakrabarti, "Monte carlo study of polymer chains end-grafted onto a spherical interface," *Physical Review E*, vol. 47, no. 6, pp. 4240–4246, Jun. 1993.

[50] T. E. Cloutier and J. Widom, "Spontaneous sharp bending of double-stranded DNA," *Molecular Cell*, vol. 14, no. 3, pp. 355–362, May 2004.

[51] P. A. Wiggins, T. v. d. Heijden, F. Moreno-Herrero, A. Spakowitz, R. Phillips, J. Widom, C. Dekker, and P. C. Nelson, "High flexibility of DNA on short length scales probed by atomic force microscopy," *Nature Nanotechnology*, vol. 1, no. 2, pp. 137–141, Nov. 2006.

[52] L. Han, H. G. Garcia, S. Blumberg, K. B. Towles, J. F. Beausang, P. C. Nelson, and R. Phillips, "Concentration and length dependence of DNA looping in transcriptional regulation," *PLOS ONE*, vol. 4, no. 5, e5621, May 2009.

[53] E. A. DiMarzio, "Proper accounting of conformations of a polymer near a surface," *The Journal of Chemical Physics*, vol. 42, no. 6, pp. 2101–2106, Mar. 1965.

[54] E. Eisenriegler, K. Kremer, and K. Binder, "Adsorption of polymer chains at surfaces: Scaling and monte carlo analyses," *The Journal of Chemical Physics*, vol. 77, no. 12, pp. 6296–6320, 1982.

[55] M. Slutsky, "Diffusion in a half-space: From lord kelvin to path integrals," *American Journal of Physics*, vol. 73, no. 4, pp. 308–314, 2005.

[56] W. Gobush, H. Yamakawa, W. H. Stockmayer, and W. S. Magee, "Statistical mechanics of wormlike chains. i. asymptotic behavior," *The Journal of Chemical Physics*, vol. 57, no. 7, pp. 2839–2843, Oct. 1972.

[57] T. D. Hahn and J. Kovac, "Computer simulation of the dynamics of a polymer chain terminally attached to a rigid flat surface," *Macromolecules*, vol. 23, no. 24, pp. 5153–5154, Nov. 1990.

[58] T. Cui, J. Ding, and J. Z. Y. Chen, "Mean first-passage times of looping of polymers with intrachain reactive monomers: Lattice monte carlo simulations," *Macromolecules*, vol. 39, no. 16, pp. 5540–5545, Aug. 2006.

[59] J. Wilhelm and E. Frey, "Radial distribution function of semiflexible polymers," *Physical Review Letters*, vol. 77, no. 12, pp. 2581–2584, Sep. 1996.

[60] L. Czapla, D. Swigon, and W. K. Olson, "Sequence-dependent effects in the cyclization of short dna," *Journal of Chemical Theory and Computation*, vol. 2, no. 3, pp. 685–695, 2006.

[61] K. B. Towles, J. F. Beausang, H. G. Garcia, R. Phillips, and P. C. Nelson, "First-principles calculation of DNA looping in tethered particle experiments," *Physical Biology*, vol. 6, no. 2, p. 025 001, Jun. 2009.

[62] N. B. Becker, A. Rosa, and R. Everaers, "The radial distribution function of wormlike chains," *The European Physical Journal E*, vol. 32, no. 1, pp. 53–69, May 2010.

[63] J. Skolnick and M. Fixman, "Electrostatic persistence length of a wormlike polyelectrolyte," *Macromolecules*, vol. 10, no. 5, 944948, 1977.

[64] T. Odijk, "Polyelectrolytes near the rod limit," *Journal of Polymer Science: Polymer Physics Edition*, vol. 15, no. 3, 477483, 1977.

[65] J. R. Wenner, M. C. Williams, I. Rouzina, and V. A. Bloomfield, "Salt dependence of the elasticity and overstretching transition of single DNA molecules," *Biophysical Journal*, vol. 82, no. 6, pp. 3160–3169, Jun. 2002.

[66] Y. Zhang and D. M. Crothers, "Statistical mechanics of sequence-dependent circular DNA and its application for DNA cyclization," *Biophysical Journal*, vol. 84, no. 1, pp. 136–153, Jan. 2003.

[67] "Definitions and nomenclature of nucleic acid structure components.," *Nucleic acids research*, vol. 17, no. 5, pp. 1797–803, Mar. 1989.

[68] A. Perez, F. Lankas, F. J. Luque, and M. Orozco, "Towards a molecular dynamics consensus view of b-dna flexibility," *Nucleic acids research*, vol. 36, no. 7, pp. 2379–2394, 2008.

[69] N. Douarche and S. Cocco, "Protein-mediated DNA loops: Effects of protein bridge size and kinks," *Physical Review E*, vol. 72, no. 6, p. 061 902, Dec. 2005.

[70] J.-F. Allemand, S Cocco, N Douarche, and G Lia, "Loops in dna: An overview of experimental and theoretical approaches," *The European Physical Journal E*, vol. 19, no. 3, pp. 293–302, 2006.

[71] H. E. Daniels, "XXI.The statistical theory of stiff chains," *Proceedings of the Royal Society of Edinburgh, Section: A Mathematics*, vol. 63, no. 03, pp. 290–311, 1952.

[72] D Thirumalai and B.-Y. Ha, "Statistical mechanics of semiflexible chains: A mean-field variational approach," *ArXiv preprint cond-mat/9705200*, 1997.

[73] R. G. Winkler, "Deformation of semiflexible chains," *The Journal of chemical physics*, vol. 118, no. 6, pp. 2919–2928, 2003.

[74] S. Stepanow and G. M. Schütz, "The distribution function of a semiflexible polymer and random walks with constraints," *EPL (Europhysics Letters)*, vol. 60, no. 4, p. 546, Nov. 2002.

[75] S. Mehraeen, B. Sudhanshu, E. F. Koslover, and A. J. Spakowitz, "End-to-end distribution for a wormlike chain in arbitrary dimensions," *Physical Review E*, vol. 77, no. 6, p. 061 803, 2008.

[76] J. Kiefer, "Sequential minimax search for a maximum," *Proceedings of the American mathematical society*, vol. 4, no. 3, pp. 502–506, 1953.

[77] W. H. Press, S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery, "Golden section search in one dimension," *Numerical Recipies in C: The Art of Scientific Computing*, vol. 2, 1992.

[78] A. S. Petrov, S. S. Douglas, and S. C. Harvey, "Effects of pulling forces, osmotic pressure, condensing agents and viscosity on the thermodynamics and kinetics of dna ejection from bacteriophages to bacterial cells: A computational study," *Journal of Physics: Condensed Matter*, vol. 25, no. 11, p. 115 101, 2013.

[79] J. Ambia-Garrido, A. Vainrub, and B. M. Pettitt, "A model for structure and thermodynamics of ssDNA and dsDNA near a surface: A coarse grained approach," *Computer Physics Communications*, vol. 181, no. 12, pp. 2001–2007, Dec. 2010.

[80] L. C. Zanetti-Domingues, M. L. Martin-Fernandez, S. R. Needham, D. J. Rolfe, and D. T. Clarke, "A systematic investigation of differential effects of cell culture substrates on the extent of artifacts in single-molecule tracking," *PLOS ONE*, vol. 7, no. 9, e45655, Sep. 2012.

[81] F. Fang, J. Satulovsky, and I. Szleifer, "Kinetics of protein adsorption and desorption on surfaces with grafted polymers," *Biophysical Journal*, vol. 89, no. 3, pp. 1516–1533, Sep. 2005.

[82] I. Szleifer, "Polymers and proteins: Interactions at interfaces," *Current Opinion in Solid State and Materials Science*, vol. 2, no. 3, pp. 337–344, 1997.

[83] P. De Gennes, "Conformations of polymers attached to an interface," *Macromolecules*, vol. 13, no. 5, pp. 1069–1075, 1980.

[84] S. T. Milner, T. A. Witten, and M. E. Cates, "Theory of the grafted polymer brush," *Macromolecules*, vol. 21, no. 8, pp. 2610–2619, Aug. 1988.

[85] J. U. Kim and M. W. Matsen, "Repulsion exerted on a spherical particle by a polymer brush," *Macromolecules*, vol. 41, no. 1, pp. 246–252, Jan. 2008.

[86] P. J. Park and W. Sung, "Polymer release out of a spherical vesicle through a pore," *Physical Review E*, vol. 57, no. 1, pp. 730–734, Jan. 1998.

[87] D. Egecioglu and J. H. Brickner, "Gene positioning and expression," *Current Opinion in Cell Biology*, vol. 23, no. 3, pp. 338–345, Jun. 2011.

[88] E. M. Green, Y. Jiang, R. Joyner, and K. Weis, "A negative feedback loop at the nuclear periphery regulates GAL gene expression," *Molecular Biology of the Cell*, vol. 23, no. 7, pp. 1367–1375, Apr. 2012.

[89] C. Hiergeist and R. Lipowsky, "Elastic properties of polymer-decorated membranes," *Journal de Physique II*, vol. 6, no. 10, 14651481, 1996.

[90] Y. W. Kim and W. Sung, "Membrane curvature induced by polymer adsorption," *Physical Review E*, vol. 63, no. 4, p. 041 910, Mar. 2001.

[91] J. Shin and W. Sung, "Effects of static and temporally fluctuating tensions on semiflexible polymer looping," *The Journal of Chemical Physics*, vol. 136, no. 4, pp. 045 101–045101–6, Jan. 2012.

[92] W. K. Kim and W. Sung, "How a single stretched polymer responds coherently to a minute oscillation in fluctuating environments: An entropic stochastic resonance," *The Journal of Chemical Physics*, vol. 137, no. 7, pp. 074 903–074903–8, Aug. 2012.

[93] R. M. Neumann, "The entropy of a single gaussian macromolecule in a noninteracting solvent," *The Journal of Chemical Physics*, vol. 66, no. 2, pp. 870–871, 1977.

[94] M. Kardar and R. Golestanian, "The frictioni of vacuum, and other fluctuation-induced forces," *Reviews of Modern Physics*, vol. 71, no. 4, p. 1233, 1999.

[95] D. Leckband and J. Israelachvili, "Intermolecular forces in biology," *Quarterly reviews of biophysics*, vol. 34, no. 02, pp. 105–267, 2001.

[96] D. E. Kharzeev, "Deconfinement as an entropic self-destruction: A solution for the quarkonium suppression puzzle?" *Physical Review D*, vol. 90, no. 7, p. 074 007, 2014.

[97] T. Jacobson, "Thermodynamics of spacetime: The einstein equation of state," *Physical Review Letters*, vol. 75, no. 7, p. 1260, 1995.

[98] E. Verlinde, "On the origin of gravity and the laws of newton," *Journal of High Energy Physics*, vol. 2011, no. 4, pp. 1–27, 2011.

[99] R. M. Neumann, "Entropic approach to brownian movement," *American Journal of Physics*, vol. 48, no. 5, pp. 354–357, 1980.

[100] D. Bartolo, A. Ajdari, J.-B. Fournier, and R. Golestanian, "Fluctuations of fluctuation-induced casimir-like forces," *Physical review letters*, vol. 89, no. 23, p. 230 601, 2002.

[101] R. M. Neumann, "Implications of using the entropy spring model for an ideal polymer chain," *Physical Review A*, vol. 34, no. 4, p. 3486, 1986.

[102] J. Weiner, "Entropic versus kinetic viewpoints in rubber elasticity," *Am. J. Phys*, vol. 55, no. 8, p. 8, 1987.

[103] U. M. B. Marconi, A. Puglisi, L. Rondoni, and A. Vulpiani, "Fluctuation–dissipation: Response theory in statistical physics," *Physics reports*, vol. 461, no. 4, pp. 111–195, 2008.

[104] Y. G. Rudoi and A. D. Sukhanov, "Thermodynamic fluctuations within the gibbs and einstein approaches," *Physics-Uspekhi*, vol. 43, no. 12, pp. 1169–1199, 2000.

[105] U. Gerland, R. Bundschuh, and T. Hwa, "Force-induced denaturation of rna," *Biophysical Journal*, vol. 81, no. 3, pp. 1324–1332, 2001.

[106] M. Manosas and F. Ritort, "Thermodynamic and kinetic aspects of rna pulling experiments," *Biophysical journal*, vol. 88, no. 5, pp. 3224–3242, 2005.

[107] U. Gerland, R. Bundschuh, and T. Hwa, "Mechanically probing the folding pathway of single rna molecules," *Biophysical journal*, vol. 84, no. 5, pp. 2831–2840, 2003.

[108] S. Turner, M Cabodi, and H. Craighead, "Confinement-induced entropic recoil of single dna molecules in a nanofluidic structure," *Physical Review Letters*, vol. 88, no. 12, p. 128 103, 2002.

[109] P. Prinsen, L. T. Fang, A. M. Yoffe, C. M. Knobler, and W. M. Gelbart, "The force acting on a polymer partially confined in a tube," *The Journal of Physical Chemistry B*, vol. 113, no. 12, pp. 3873–3879, 2009.

[110] S.-L. Zhao, J. Wu, D. Gao, and J. Wu, "Gaussian fluctuations in tethered dna chains," *The Journal of chemical physics*, vol. 134, no. 6, p. 065 103, 2011.

[111] K. Guo, F. Qiu, H. Zhang, and Y. Yang, "Rigid rod anchored to infinite membrane," *The Journal of chemical physics*, vol. 123, no. 7, p. 074 906, 2005.

[112] J. Odenheimer, D. W. Heermann, and M. Brill, "Forces by and on a polymer grafted to a repulsive wall," *International Journal of Modern Physics C*, vol. 16, no. 10, pp. 1561–1575, 2005.

[113] T Bickel, C Jeppesen, and C. Marques, "Local entropic effects of polymers grafted to soft interfaces," *The European Physical Journal E*, vol. 4, no. 1, pp. 33–43, 2001.

[114] C. Jeppesen, J. Y. Wong, T. L. Kuhl, J. N. Israelachvili, N. Mullah, S. Zalipsky, and C. M. Marques, "Impact of polymer tether length on multiple ligand-receptor bond formation," *Science*, vol. 293, no. 5529, pp. 465–468, 2001.

[115] Z. Guo, M. Gibson, S. Sitha, S. Chu, and U. Mohanty, "Role of large thermal fluctuations and magnesium ions in t-rna selectivity of the ribosome," *Proceedings of the National Academy of Sciences*, vol. 108, no. 10, pp. 3947–3951, 2011.

[116] N. Van Kampen and J. Lodder, "Constraints," *American Journal of Physics*, vol. 52, no. 5, pp. 419–424, 1984.

[117] M. Fixman, "Classical statistical mechanics of constraints: A theorem and application to polymers," *Proceedings of the National Academy of Sciences*, vol. 71, no. 8, pp. 3050–3053, 1974.

[118] J. Schlitter and M. Klähn, "The free energy of a reaction coordinate at multiple constraints: A concise formulation," *Molecular Physics*, vol. 101, no. 23-24, pp. 3439–3443, 2003.

[119] J. T. Waters and H. D. Kim, "Equilibrium statistics of a surface-pinned semiflexible polymer," *Macromolecules*, vol. 46, no. 16, pp. 6659–6666, 2013.

[120] W. B. Ludington, H. Ishikawa, Y. V. Serebrenik, A. Ritter, R. A. Hernandez-Lopez, J. Gunzenhauser, E. Kannegaard, and W. F. Marshall, "A systematic comparison of mathematical models for inherent measurement of ciliary length: How a cell can measure length and volume," *Biophysical journal*, vol. 108, no. 6, pp. 1361–1379, 2015.

[121] A Milchev, L Klushin, A Skvortsov, and K Binder, "Ejection of a polymer chain from a nanopore: Theory and computer experiment," *Macromolecules*, vol. 43, no. 16, pp. 6877–6885, 2010.

[122] H. C. Andersen, "Molecular dynamics simulations at constant pressure and/or temperature," *The Journal of chemical physics*, vol. 72, no. 4, pp. 2384–2393, 1980.

[123] S. S. Sheiko, S. Panyukov, and M. Rubinstein, "Bond tension in tethered macromolecules," *Macromolecules*, vol. 44, no. 11, pp. 4520–4529, 2011.

[124] Z. Farkas, I. Derényi, and T. Vicsek, "Dna uptake into nuclei: Numerical and analytical results," *Journal of Physics: Condensed Matter*, vol. 15, no. 18, S1767, 2003.

[125] C. F. Lee, "Thermal breakage of a semiflexible polymer: Breakage profile and rate," *Journal of Physics: Condensed Matter*, vol. 27, no. 27, p. 275 101, 2015.

[126] J. T. Waters and H. D. Kim, "Calculation of a fluctuating entropic force by phase space sampling," *Physical Review E*, vol. 92, p. 013 308, Jul. 2015.

[127] G. M. Leuty, M. Tsige, G. S. Grest, and M. Rubinstein, "Tension amplification in tethered layers of bottle-brush polymers," *Macromolecules*, vol. 49, no. 5, pp. 1950–1960, 2016.

[128] Y. Hammer and Y. Kantor, "Ideal polymers near scale-free surfaces," *Physical Review E*, vol. 89, no. 2, p. 022 601, 2014.

[129] S. I. Stoliarov, P. R. Westmoreland, M. R. Nyden, and G. P. Forney, "A reactive molecular dynamics model of thermal decomposition in polymers: I. poly (methyl methacrylate)," *Polymer*, vol. 44, no. 3, pp. 883–894, 2003.

[130] S. I. Stoliarov, R. E. Lyon, and M. R. Nyden, "A reactive molecular dynamics model of thermal decomposition in polymers. ii. polyisobutylene," *Polymer*, vol. 45, no. 25, pp. 8613–8621, 2004.

[131] V. D. Knyazev, "Effects of chain length on the rates of cc bond dissociation in linear alkanes and polyethylene," *The Journal of Physical Chemistry A*, vol. 111, no. 19, pp. 3875–3883, 2007.

[132] K. Smith, M Bruns, S. Stoliarov, M. Nyden, O. Ezekoye, and P. Westmoreland, "Assessing the effect of molecular weight on the kinetics of backbone scission reactions in polyethylene using reactive molecular dynamics," *Polymer*, vol. 52, no. 14, pp. 3104–3111, 2011.

[133] H.-J. Lin, Y.-J. Sheng, H.-Y. Chen, and H.-K. Tsao, "Influences of linkage stiffness on rupture rate in single-molecule pulling experiments," *The Journal of Physical Chemistry B*, vol. 111, no. 23, pp. 6493–6500, 2007.

[134] O. K. Wong, M. Guthold, D. A. Erie, and J. Gelles, "Interconvertible lac repressor–dna loops revealed by single-molecule experiments," *PLoS Biol*, vol. 6, no. 9, e232, 2008.

[135] T. T. Le and H. D. Kim, "Probing the elastic limit of dna bending," *Nucleic acids research*, vol. 42, no. 16, pp. 10 786–10 794, 2014.

[136] J. Lawrimore, P. A. Vasquez, M. R. Falvo, R. M. Taylor, L. Vicci, E. Yeh, M. G. Forest, and K. Bloom, "Dna loops generate intracentromere tension in mitosis," *The Journal of cell biology*, vol. 210, no. 4, pp. 553–564, 2015.

[137] B. A. Laurent and S. M. Grayson, "Synthetic approaches for the preparation of cyclic polymers," *Chemical Society Reviews*, vol. 38, no. 8, pp. 2202–2213, 2009.

[138] A. Deffieux and M. Schappacher, "Naturally occurring and synthetic cyclic macromolecules," *Cellular and molecular life sciences*, vol. 66, no. 15, pp. 2599–2602, 2009.

[139] S. L. Craig, "Mechanochemistry: A tour of force," *Nature*, vol. 487, no. 7406, pp. 176–177, 2012.

[140] M. Frank-Kamenetskii, A. Lukashin, V. Anshelevich, and A. Vologodskii, "Torsional and bending rigidity of the double helix from data on small dna rings," *Journal of Biomolecular Structure and Dynamics*, vol. 2, no. 5, pp. 1005–1012, 1985.

[141] S. K. Kumar, M. Vacatello, and D. Y. Yoon, "Off-lattice monte carlo simulations of polymer melts confined between two plates," *The Journal of chemical physics*, vol. 89, no. 8, pp. 5206–5215, 1988.

[142] M. R. Betancourt, "Efficient monte carlo trial moves for polypeptide simulations.," *The Journal of chemical physics*, vol. 123, no. 17, pp. 174 905–174 905, 2005.

[143] J. Nilmeier, L. Hua, E. A. Coutsias, and M. P. Jacobson, "Assessing protein loop flexibility by hierarchical monte carlo sampling," *Journal of chemical theory and computation*, vol. 7, no. 5, pp. 1564–1574, 2011.

[144] S. Bottaro, W. Boomsma, K. E. Johansson, C. Andreetta, T. Hamelryck, and J. Ferkinghoff-Borg, "Subtle monte carlo updates in dense molecular systems," *Journal of Chemical Theory and Computation*, vol. 8, no. 2, pp. 695–702, 2012.

[145] S. Zamuner, A. Rodriguez, F. Seno, and A. Trovato, "An efficient algorithm to perform local concerted movements of a chain molecule.," *PloS one*, vol. 10, no. 3, e0118342, 2015.

[146] F Bosi, D Misseroni, F Dal Corso, and D Bigoni, "Development of configurational forces during the injection of an elastic rod," *Extreme Mechanics Letters*, vol. 4, pp. 83–88, 2015.

[147] J. Weiner and D. Berman, "Axial force fluctuations in long-chain molecules," *Journal of Polymer Science Part B: Polymer Physics*, vol. 24, no. 2, pp. 389–394, 1986.

[148] R. Winkler and P Reineker, "Finite size distribution and partition functions of gaussian chains: Maximum entropy approach," *Macromolecules*, vol. 25, no. 25, pp. 6891–6896, 1992.

[149] A. Planes and E. Vives, "Entropic formulation of statistical mechanics," *Journal of statistical physics*, vol. 106, no. 3-4, pp. 827–850, 2002.

[150] Y Mishin, "Thermodynamic theory of equilibrium fluctuations," *Annals of Physics*, vol. 363, pp. 48–97, 2015.

[151] M. Sprik and G. Ciccotti, "Free energy from constrained molecular dynamics," *The Journal of chemical physics*, vol. 109, no. 18, pp. 7737–7744, 1998.

[152] W. Den Otter and W. Briels, "The calculation of free-energy differences by constrained molecular-dynamics simulations," *The Journal of chemical physics*, vol. 109, no. 11, pp. 4139–4146, 1998.

[153] W. K. Den Otter, "Revisiting the exact relation between potential of mean force and free-energy profile," *Journal of chemical theory and computation*, vol. 9, no. 9, pp. 3861–3865, 2013.

[154] R. Phillips, J. Kondev, J. Theriot, and H. Garcia, *Physical biology of the cell*. Garland Science, 2012.

[155] R. Kubo, *Statistical Mechanics*. North-Holland, 1990.

[156] J. Weiner and D Perchak, "Frenkel's governor model for stretched polymers," *Macromolecules*, vol. 14, no. 5, pp. 1590–1591, 1981.

[157] P Reineker and R. Winkler, "Deterministic chaos in the motion of chain molecules," in *Relaxation in Polymers*, Springer, 1989, pp. 101–109.

[158] N. Roos, "Entropic forces in brownian motion," *American Journal of Physics*, vol. 82, no. 12, pp. 1161–1166, 2014.

[159] D. Bertolini, G. Cinacchi, and A. Tani, "On the distribution functions of depletion interactions," *The Journal of Physical Chemistry B*, vol. 115, no. 20, pp. 6608–6615, 2011.

[160] E. Villa, A. Balaeff, and K. Schulten, "Structural dynamics of the lac repressor–dna complex revealed by a multiscale simulation," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 102, no. 19, pp. 6783–6788, 2005.

[161] S. Chen and G. D. Doolen, "Lattice boltzmann method for fluid flows," *Annual review of fluid mechanics*, vol. 30, no. 1, pp. 329–364, 1998.